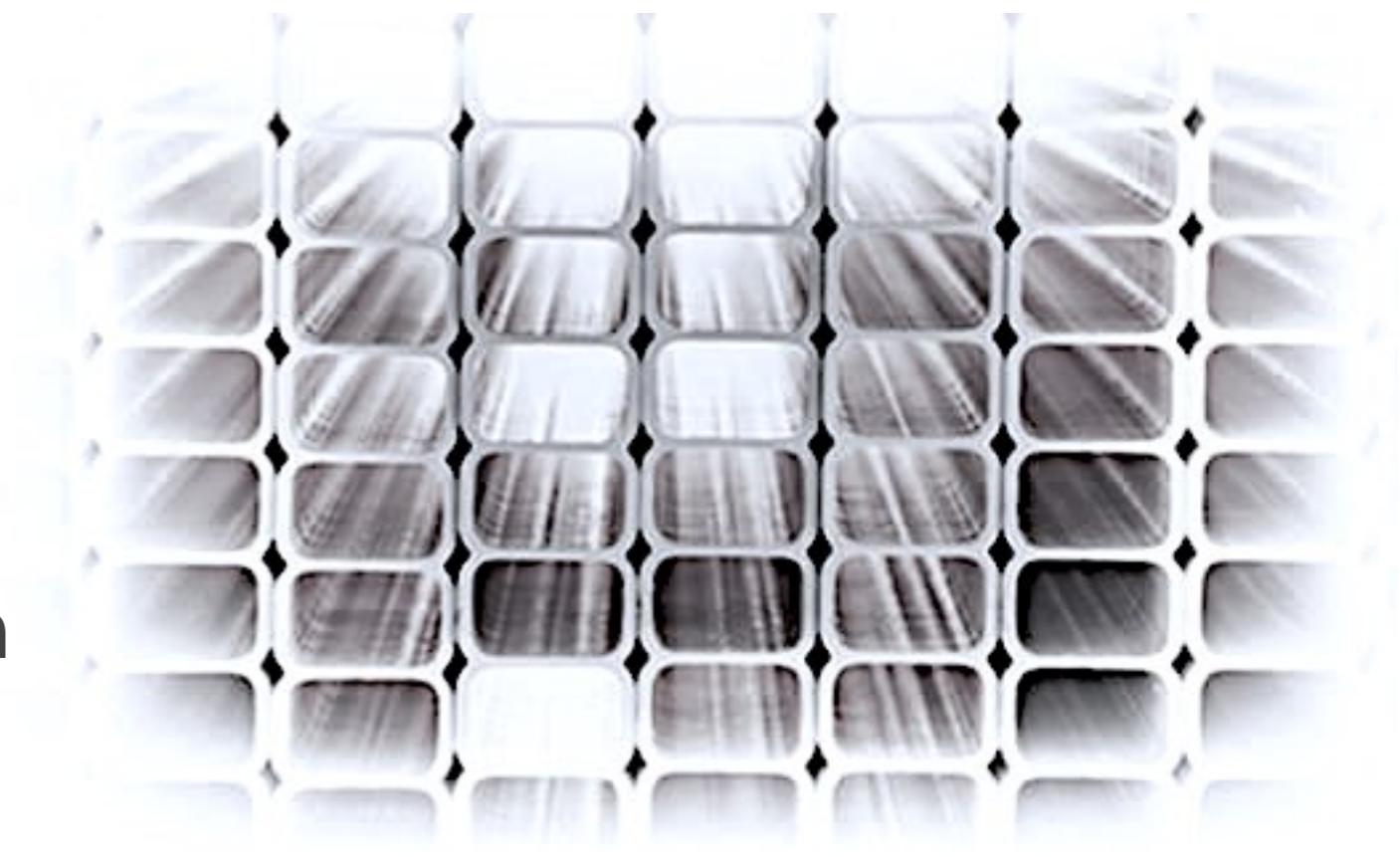
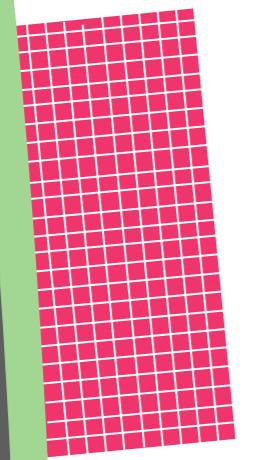


DEEP LEARNING APPLICATIONS ON NOvA

*a summary of implementations for identification
and reconstruction of NOvA events*

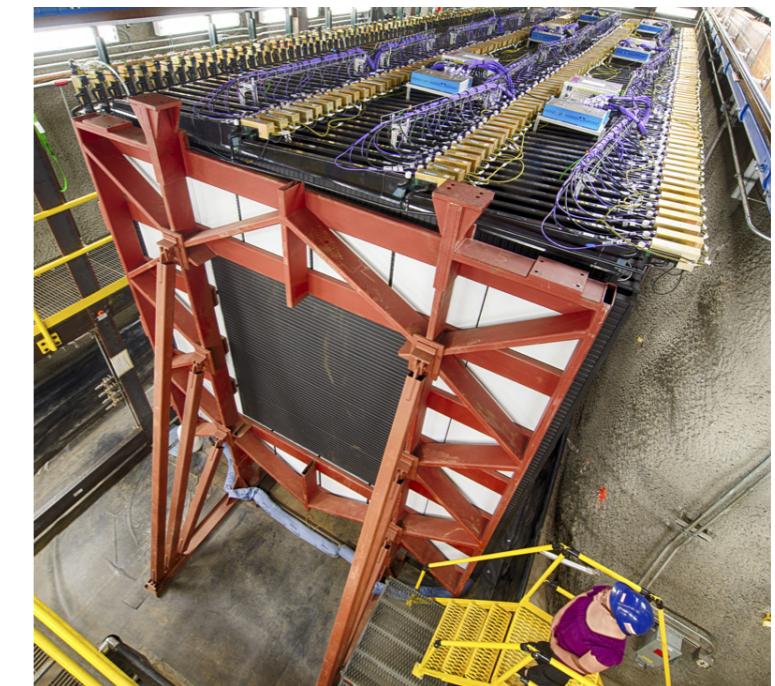
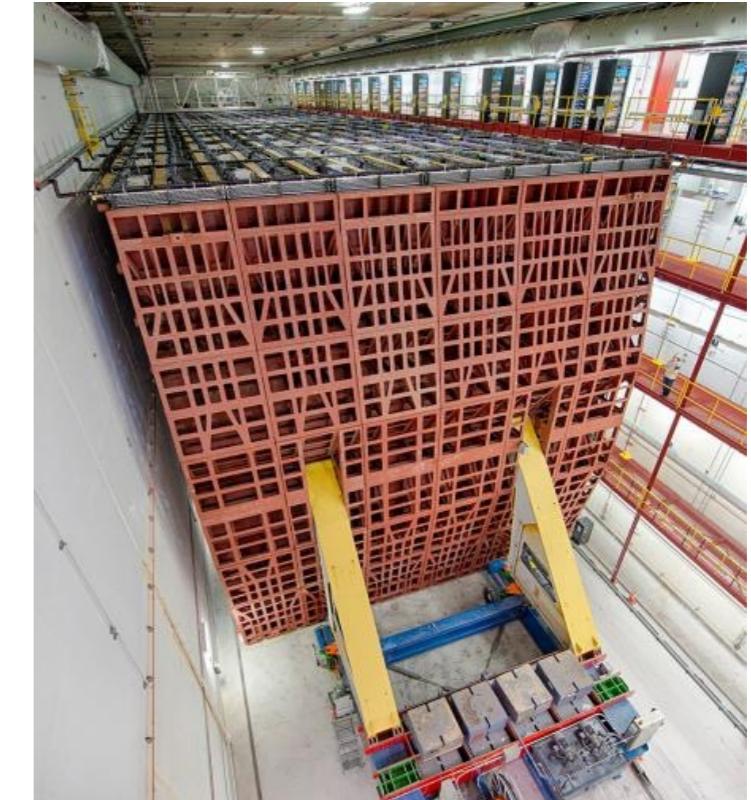
Fernanda Psihas
Ψ Indiana University
 For the NOvA Collaboration





The NOvA Experiment

Technology: A beam of ν_μ and two functionally equivalent calorimeters.





The NOvA Experiment

Technology: A beam of ν_μ and two functionally equivalent calorimeters.

Analyses:

Neutrino Oscillations

Steriles & NSI Searches

Cross Sections

Supernova Neutrinos,

Cosmic Ray Physics

Exotics Searches.

Event Reconstruction in the NOvA *Mr. Biswaranjan BEHERA*
Experiment

Comitium, Fermi National Accelerator Laboratory 13:48 - 14:06

Energy reconstruction of NOvA *Ms. Fernanda PSIHAS*
neutrino events.

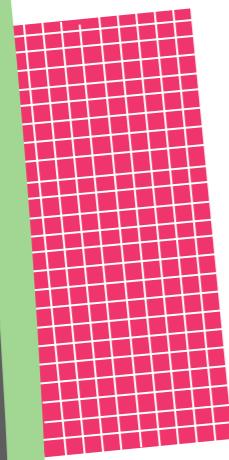
Comitium, Fermi National Accelerator Laboratory 14:06 - 14:24

Extracting neutrino oscillation parameters using a simultaneous fit of the nue appearance and *Mr. Prabhjot Singh SINGH*
numu disappearance data in the NOvA experiment

Ramsey Auditorium, Fermi National Accelerator Laboratory 10:45 - 11:03

Physics reach of electron neutrino appearance measurements in NOvA *Erika CATANO MUR*

Ramsey Auditorium, Fermi National Accelerator Laboratory 11:03 - 11:21



The NOvA Experiment

Technology: A beam of ν_μ and two functionally equivalent calorimeters.

Analyses:

Neutrino Oscillations

Steriles & NSI Searches

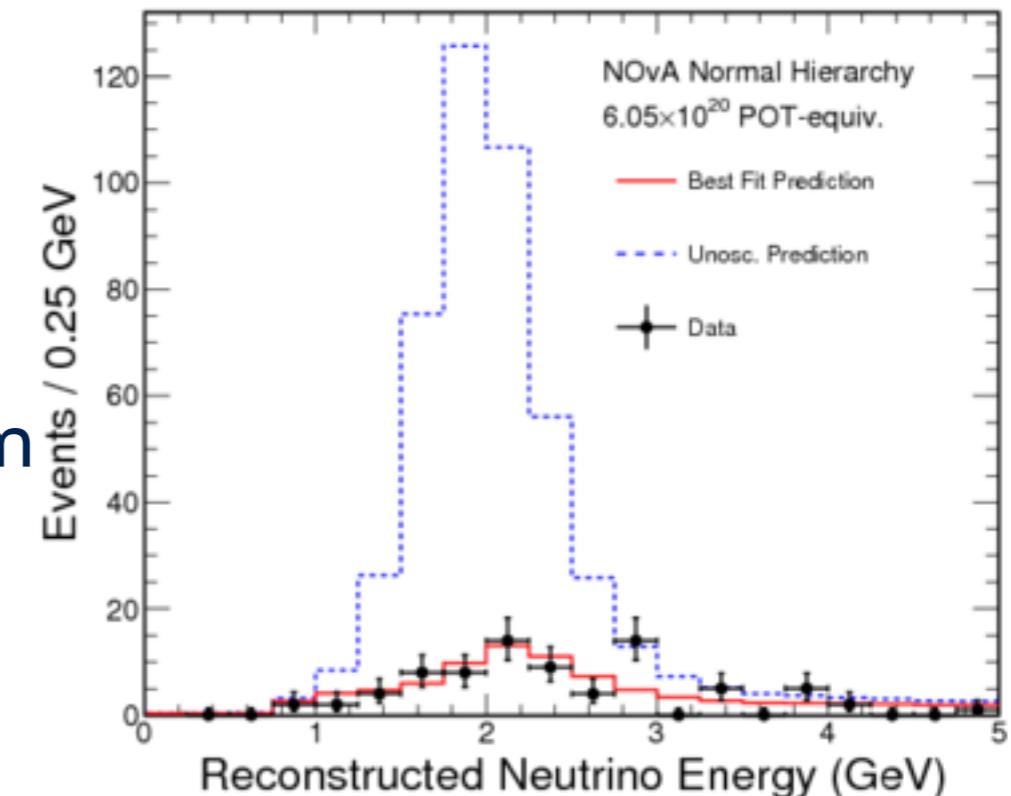
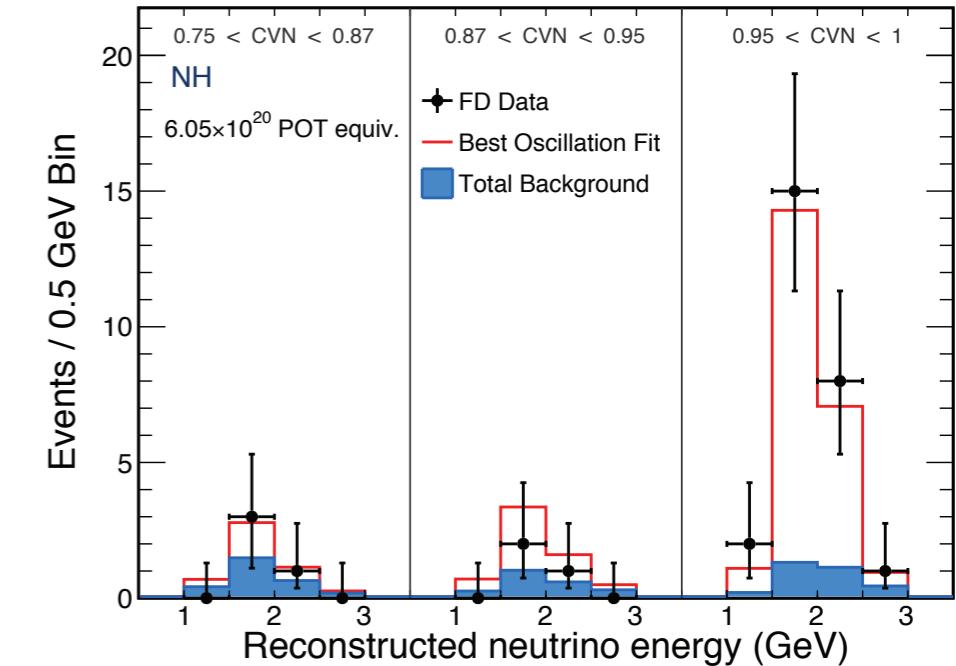
Cross Sections

Supernova Neutrinos,

Cosmic Ray Physics

Exotics Searches.

Measurable: An energy spectrum of each neutrino flavor.

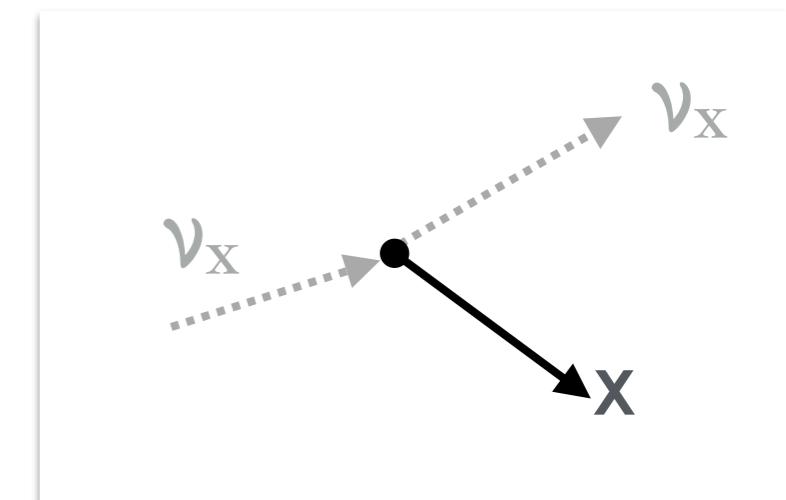


Flavor Identification

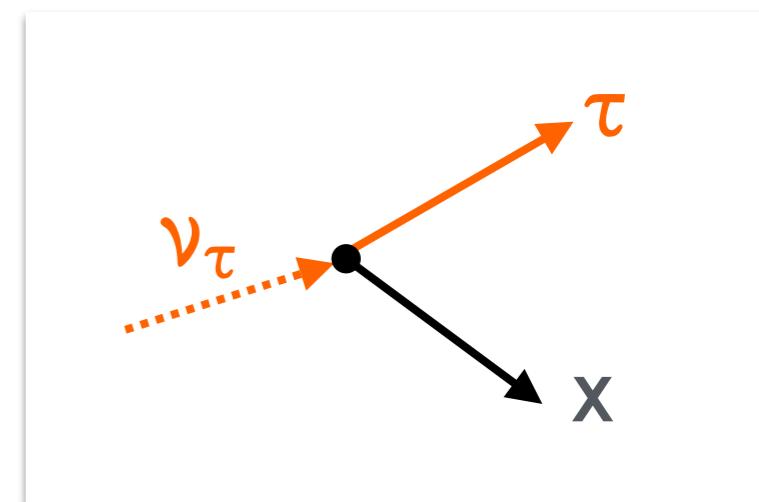
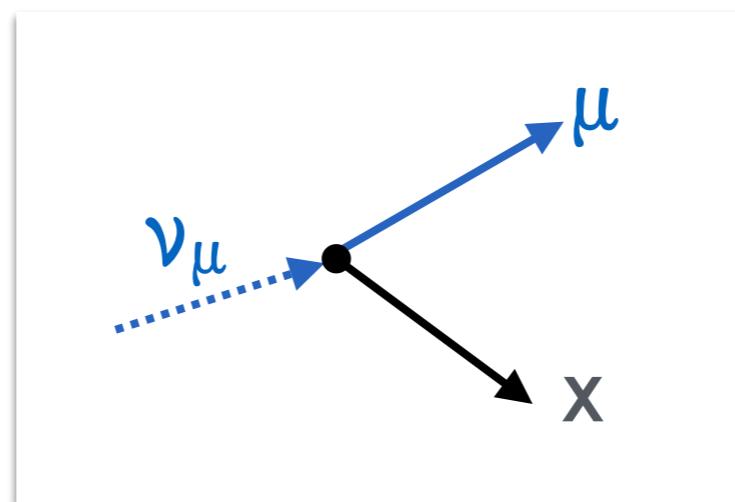
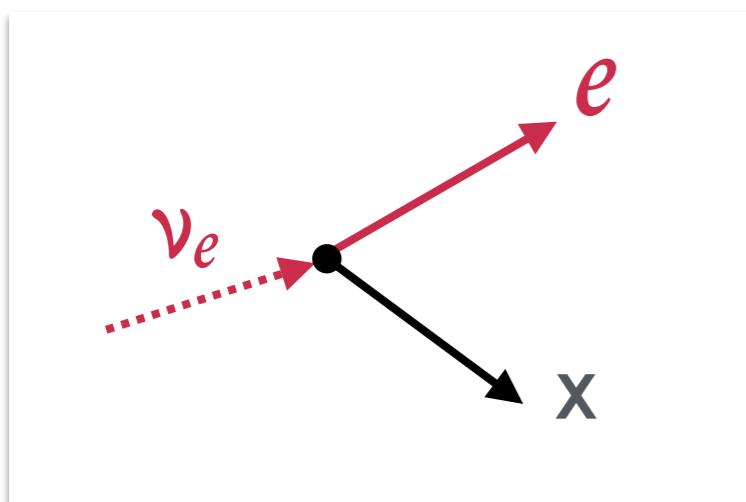
All analyses rely on identification and reconstruction.

Neutrino interactions are flavor conserving, thus, they can be identified from the outgoing lepton.

Unless, of course, it is also a neutrino.



NEUTRAL CURRENT INTERACTIONS



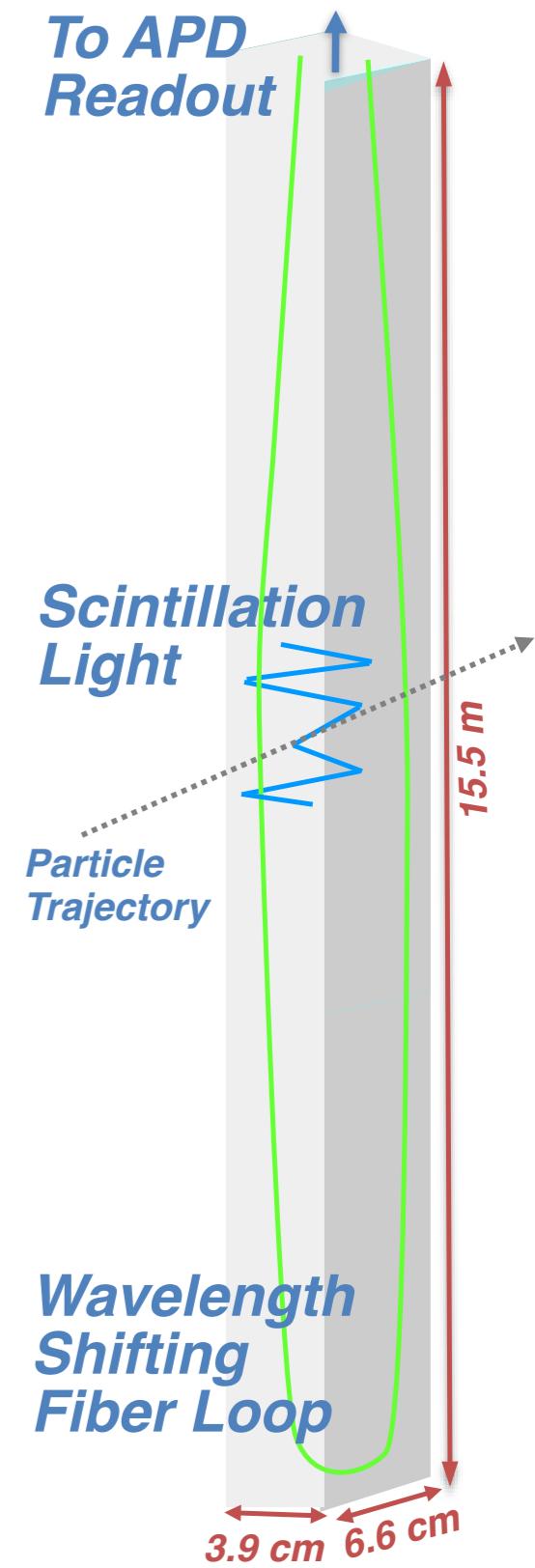
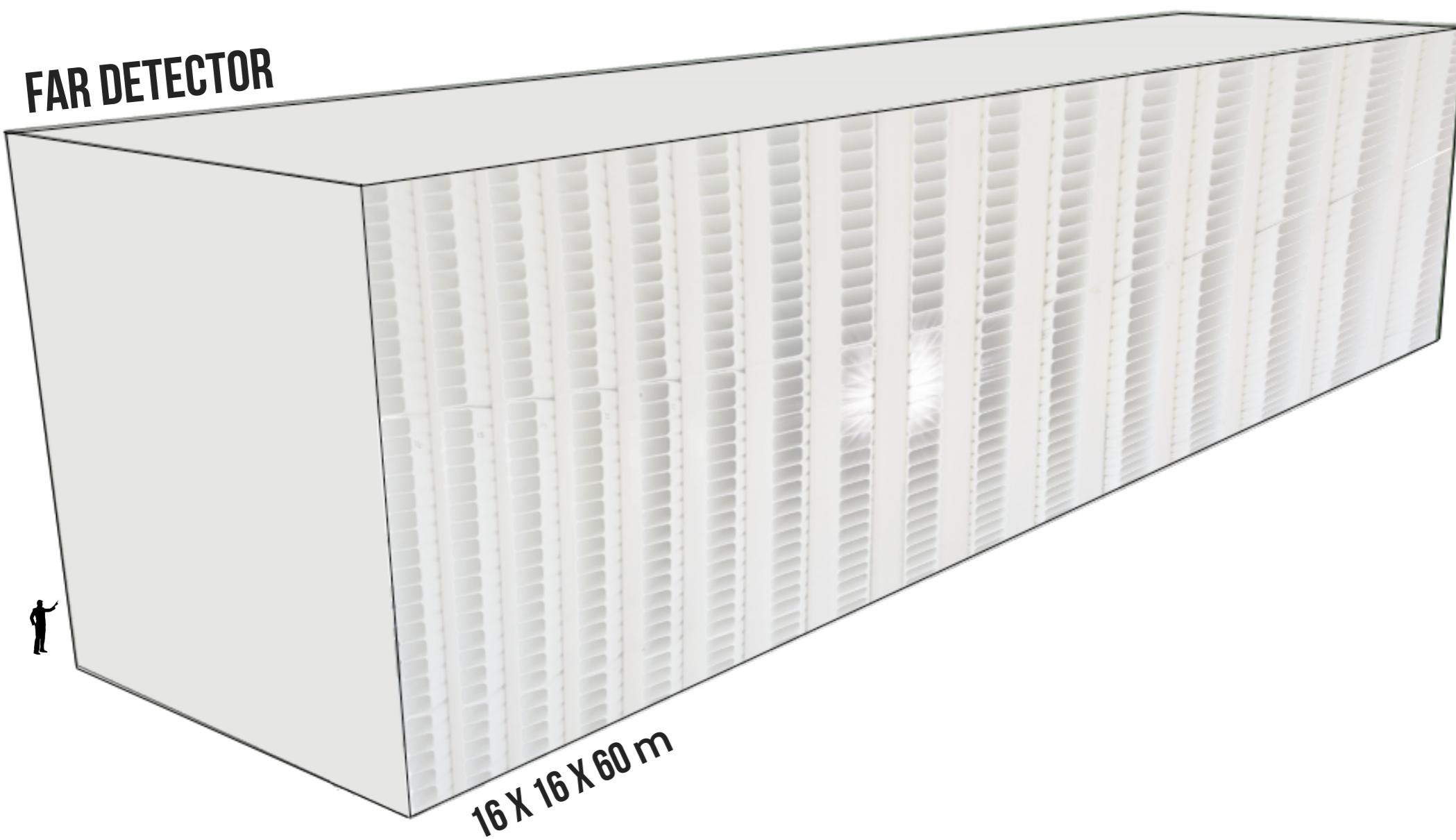
CHARGED CURRENT INTERACTIONS

The NO_νA Detectors

Sampling calorimeters optimized for electron identification.

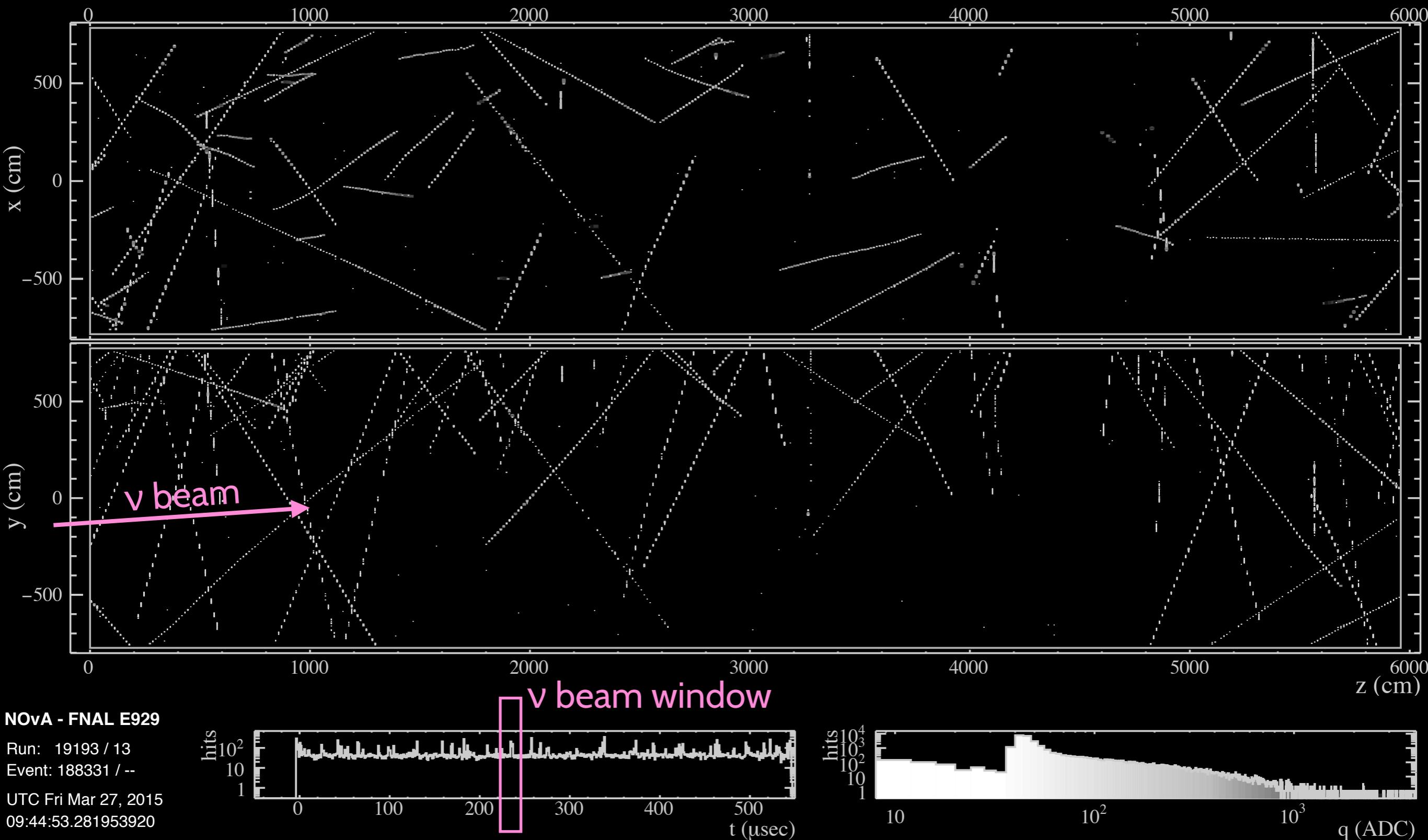
Detection of scintillation light.

Large volume at the Far detector to maximize signal statistics.

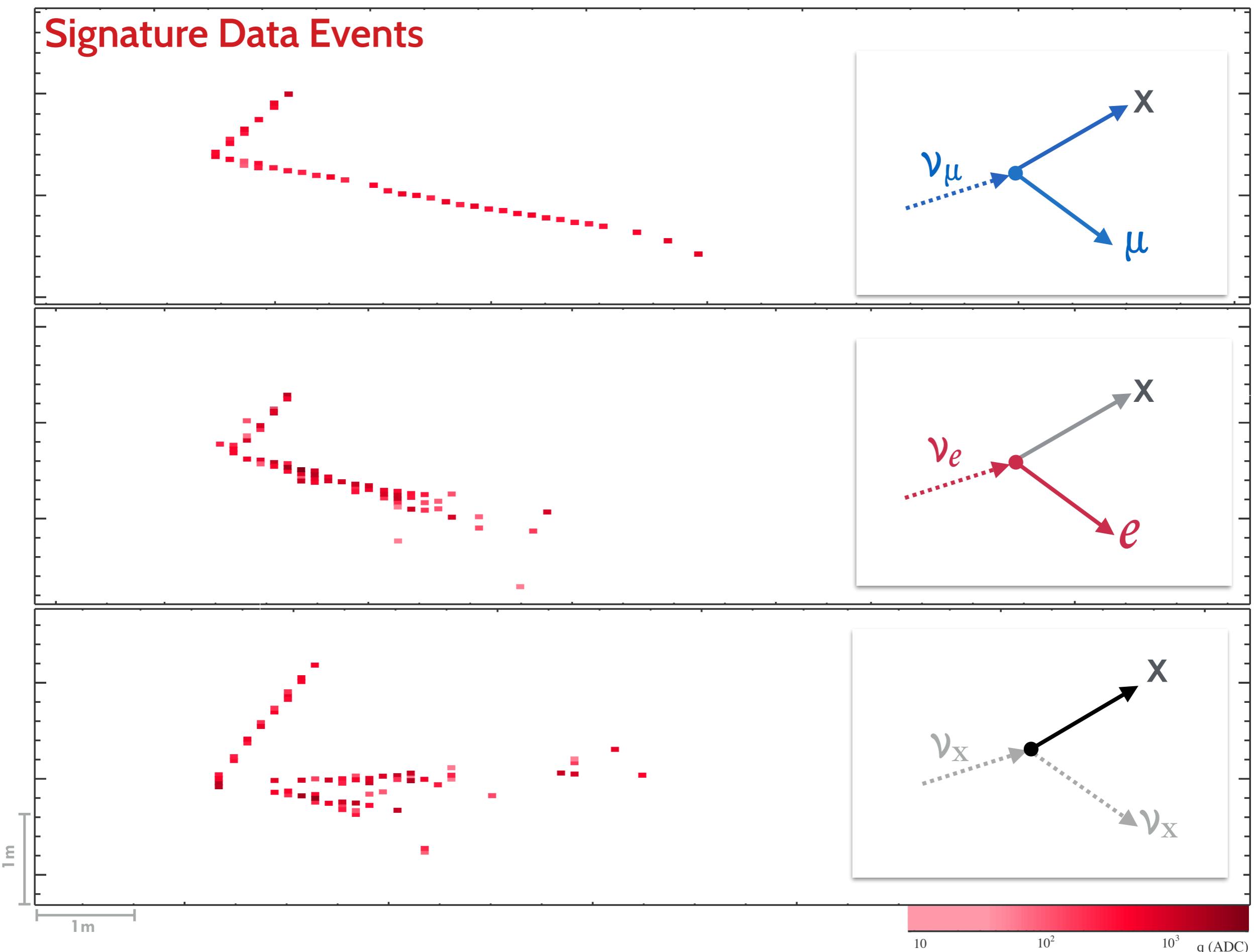


NOvA Readout and Neutrino Interactions

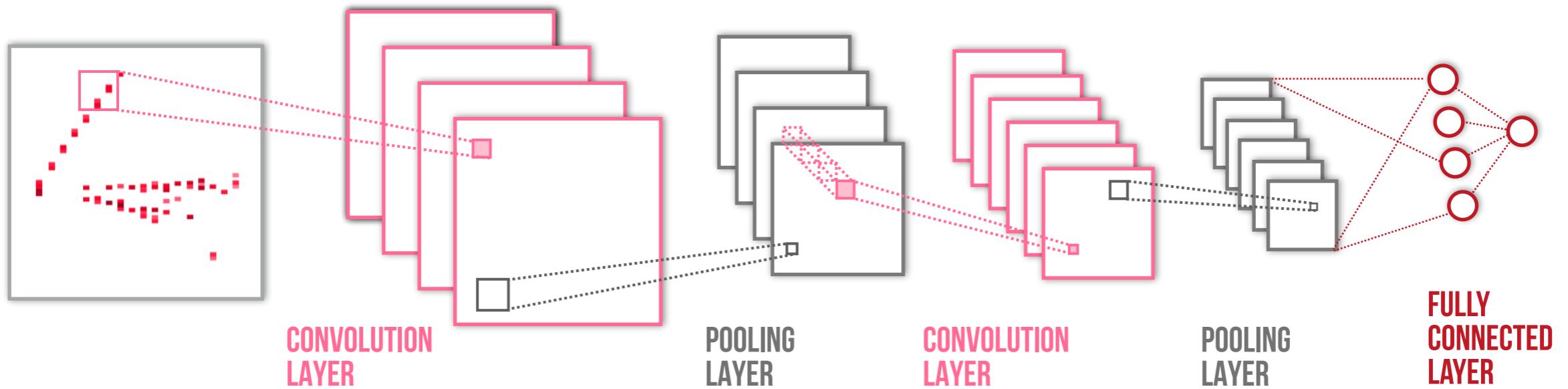
Events are 550 μs readouts around the neutrino beam spill.



Signature Data Events



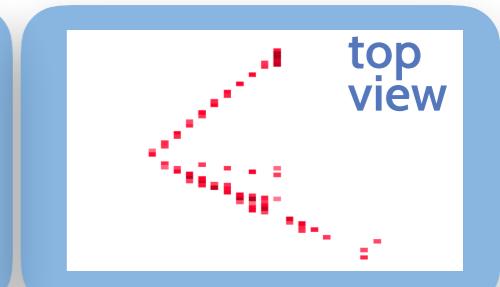
Convolutional Neural Networks



Premise: Allow the network to extract features rather than selecting them a-priori.

In Practice: Cast calibrated detector signals into maps and use CNNs to classify in the style of image recognition.

CNNs for NOvA Events

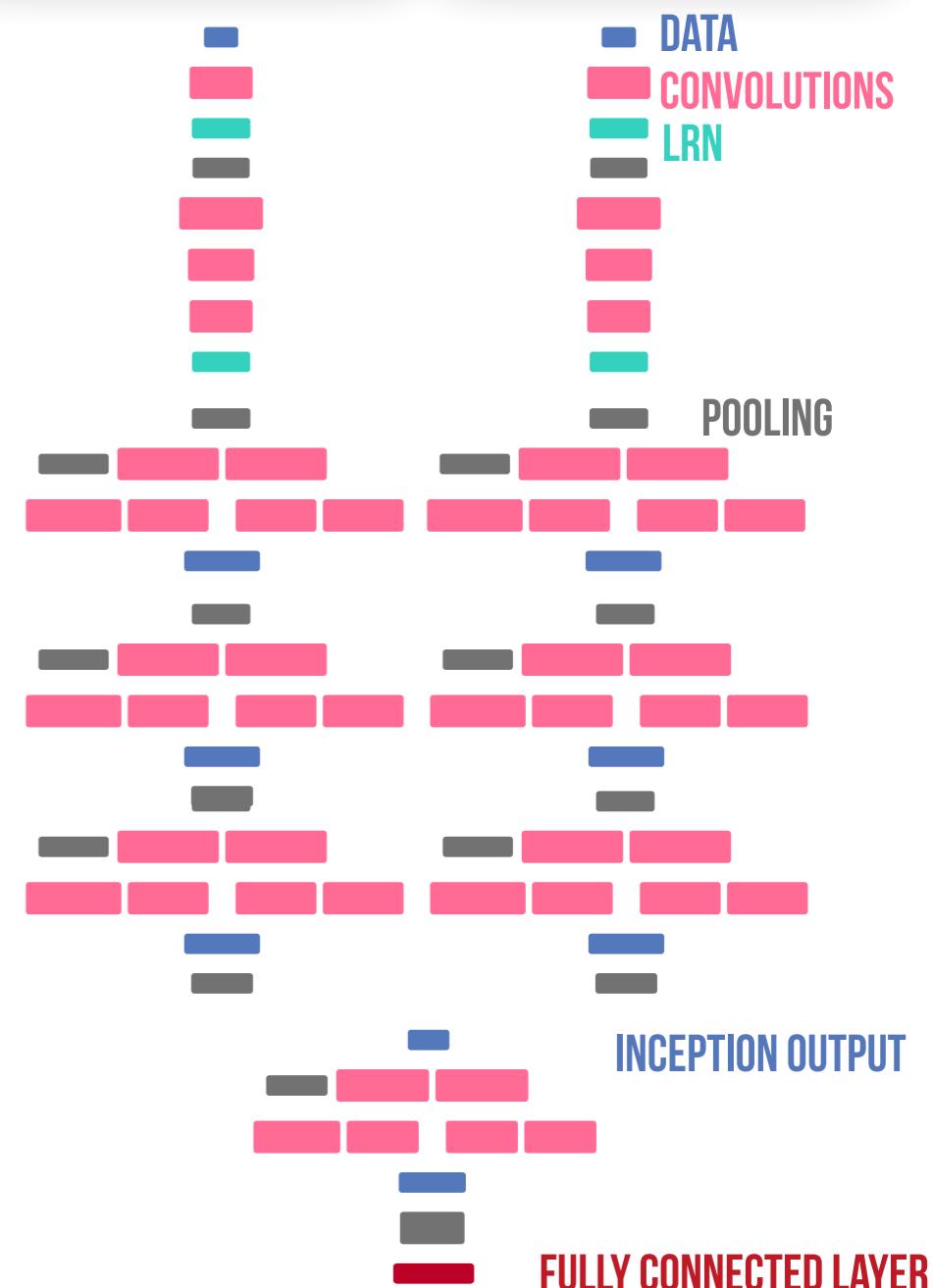


Inspired by GoogLeNet and other CNN architectures.

Trained on 4.7 million MC + cosmic ray data events

Optimized for overall accuracy and analysis FOM.

- ★ It **effectively increased our exposure by 30%** compared to traditional ID methods.
- ★ NOvA's nue appearance analysis is the **first implementation of a CNN in a HEP result.**
- ★ The CVN implementation has been extended to work on LAr-TPC detector technology.

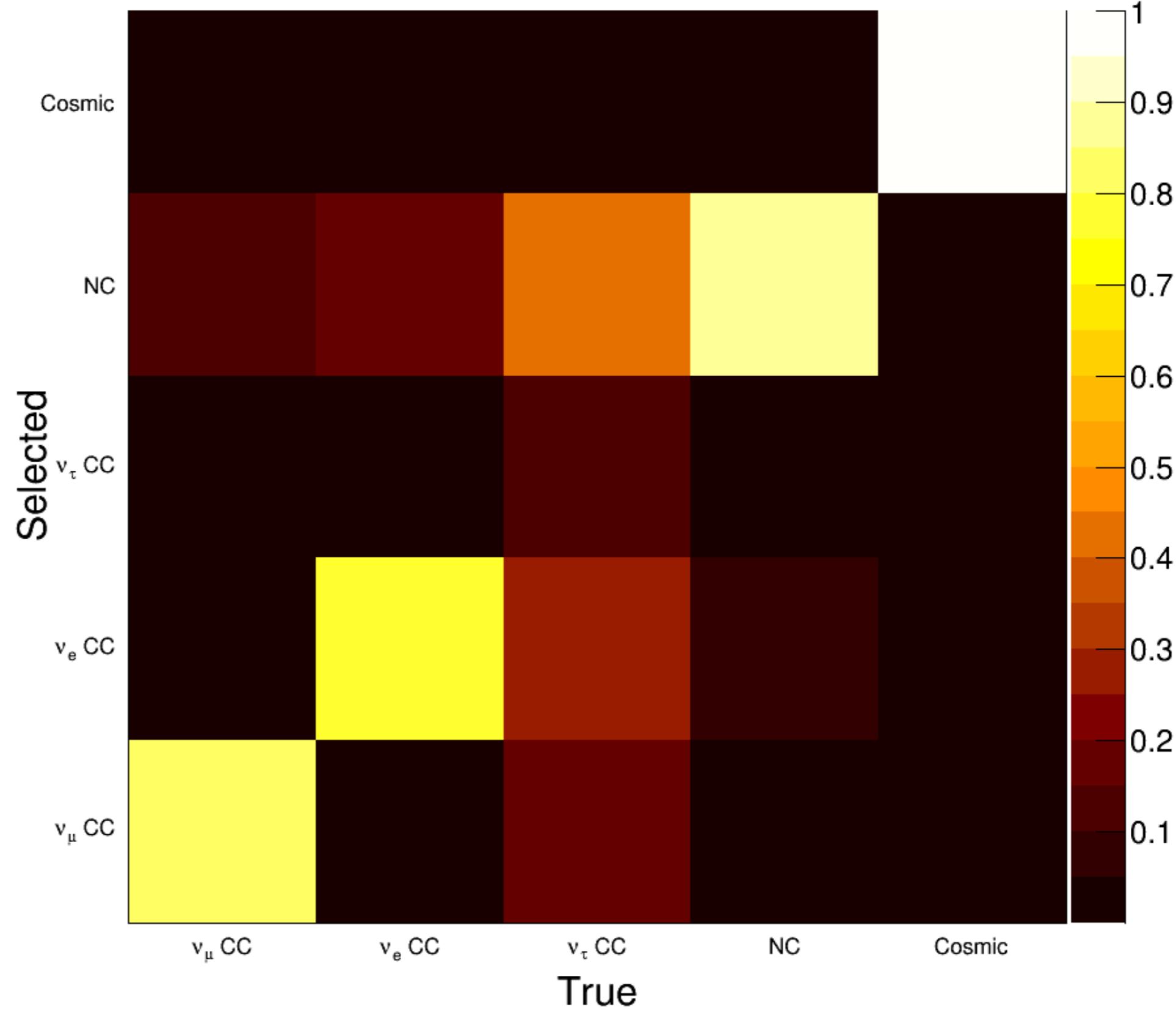
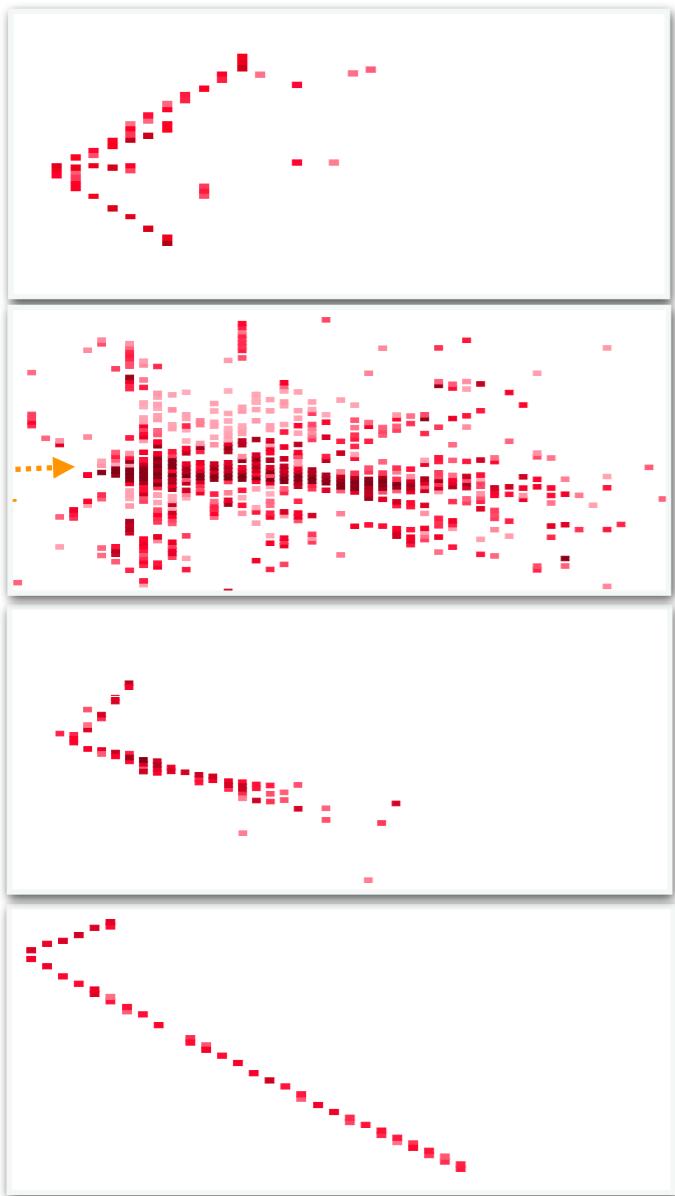


*CVN Paper: "A Convolutional Neural Network Neutrino Event Classifier"
A.Aurisano et. al. JINST 11 (2016) no.09, P09001*

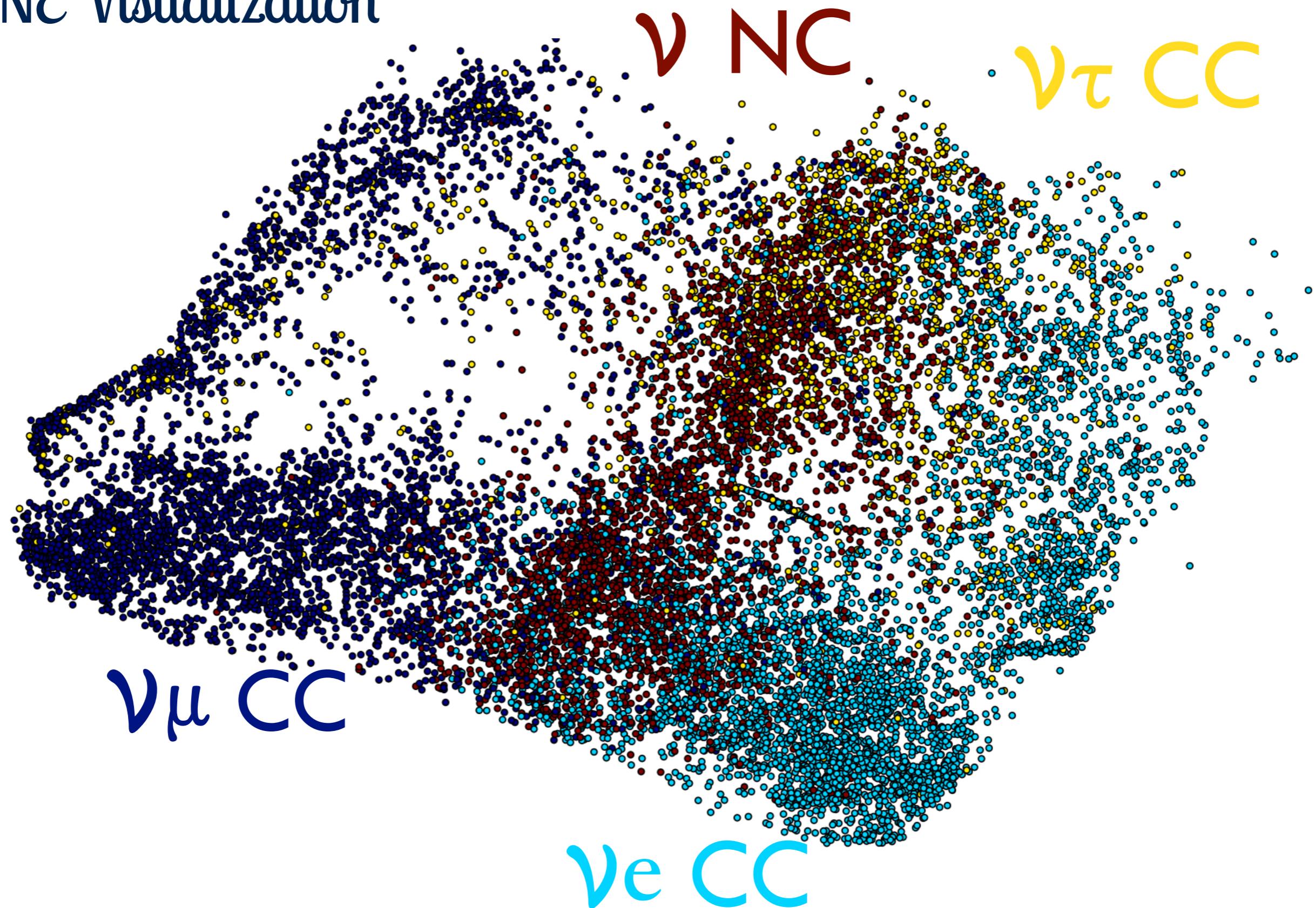


CVN Classification Matrix

NOvA Simulation

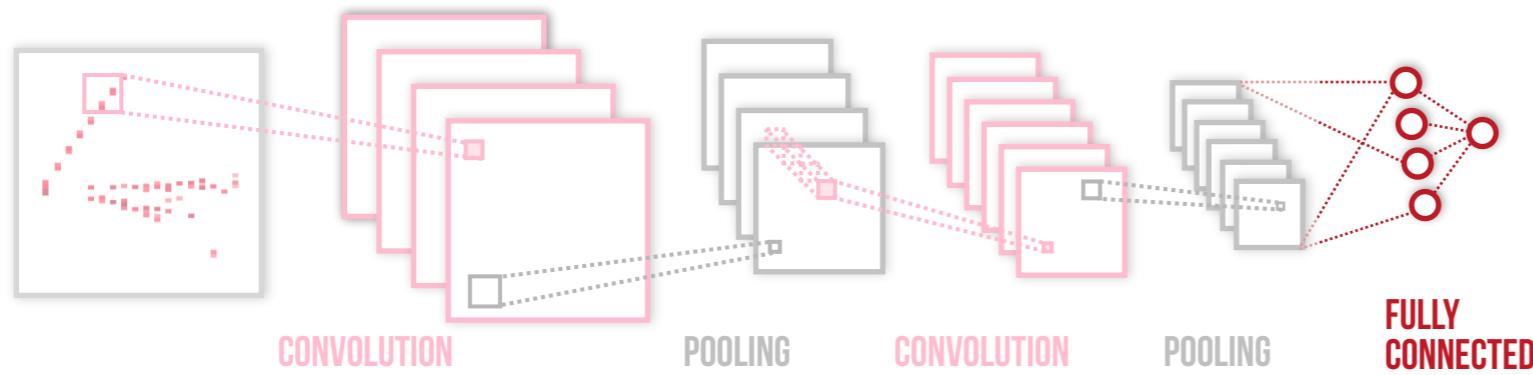


t-SNE Visualization

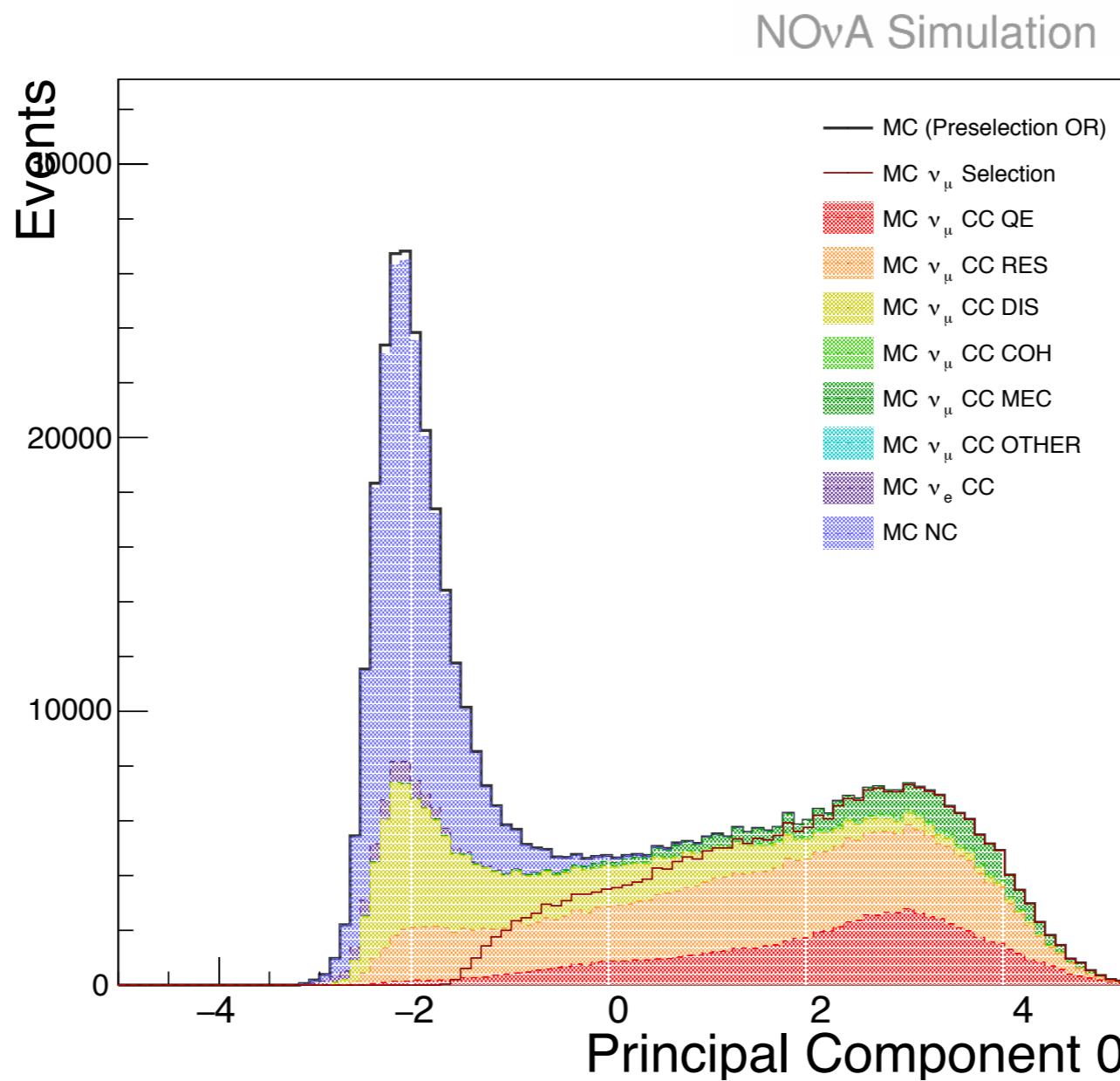


<https://indico.io/blog/visualizing-with-t-sne/>

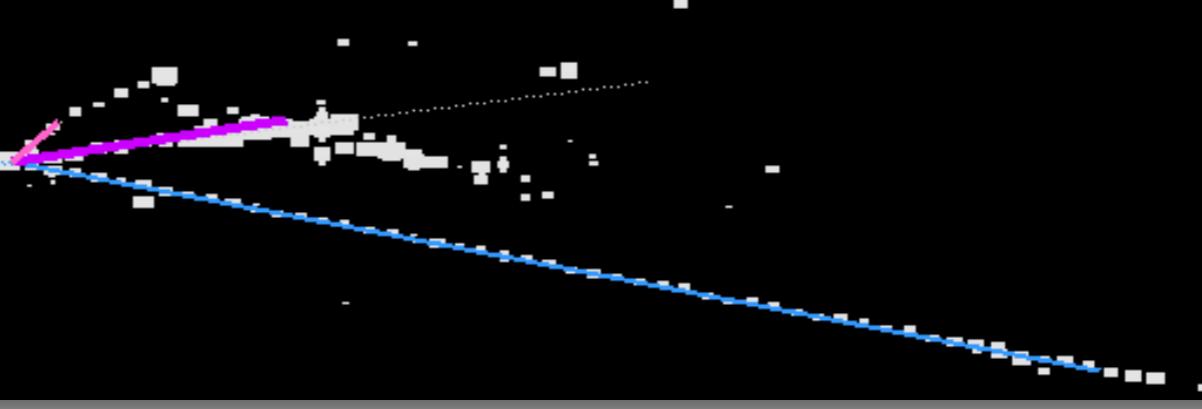
CVN Feature Principal Component Analysis



Link principal components to physical quantities and use to train simpler networks.



Deep Learning Reconstruction Program



Particle by particle tagging

Energy Reconstruction

Vertex reconstruction

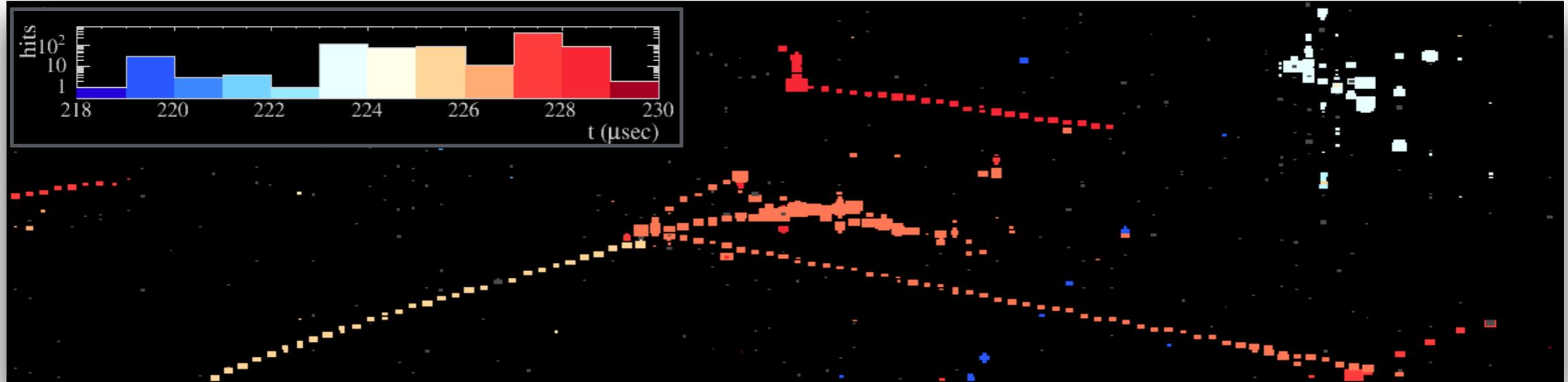
Full event identification hit by hit



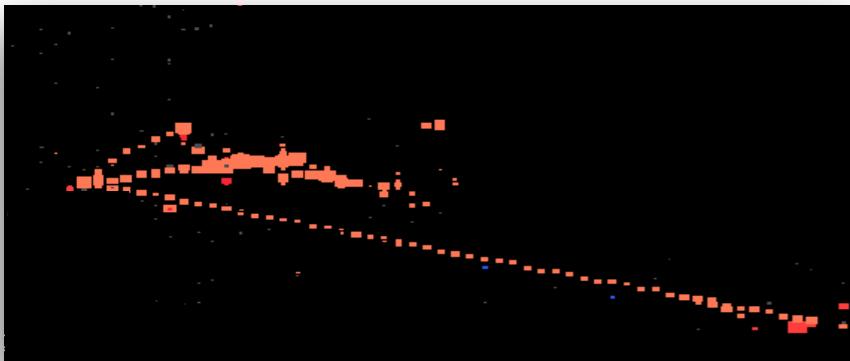
Traditional Reconstruction

Use the topology and magnitude of the energy depositions.

Takes advantage of the granularity and time resolution of our detectors.

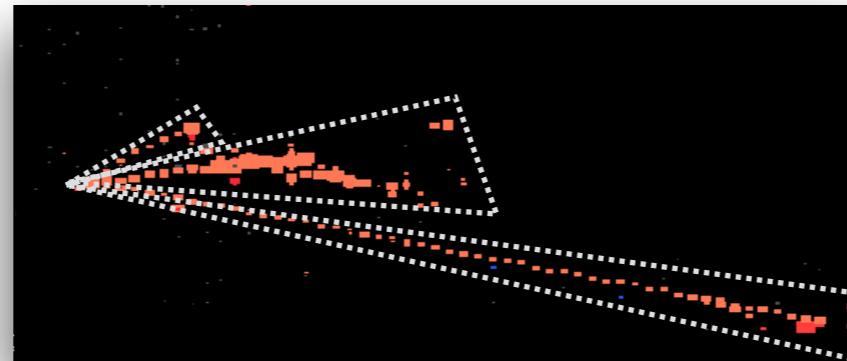


ISOLATE THE EVENT



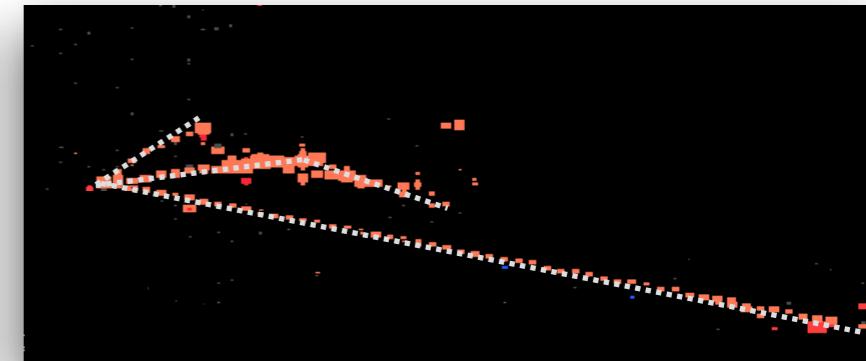
We isolate individual interactions using time and space correlation of the hits.

DEFINE CLUSTERS



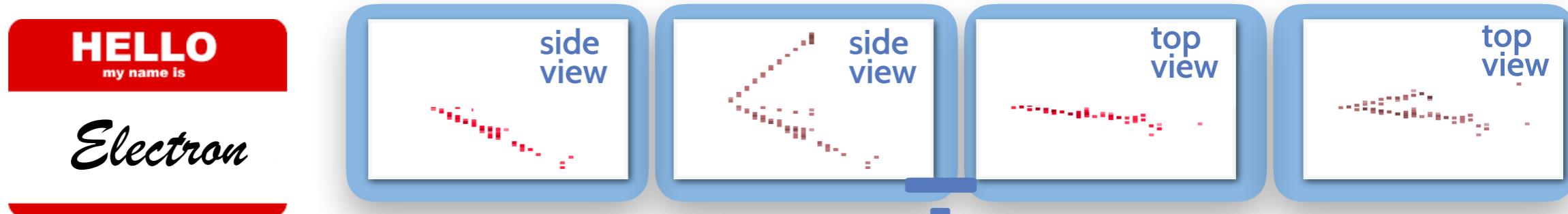
Groups of hits can be clustered as following the path of same particle starting at the interaction point.

FIT TRAJECTORY



When necessary we can fit an assumed trajectory for each cluster of hits.

CNNs for Particle Tagging

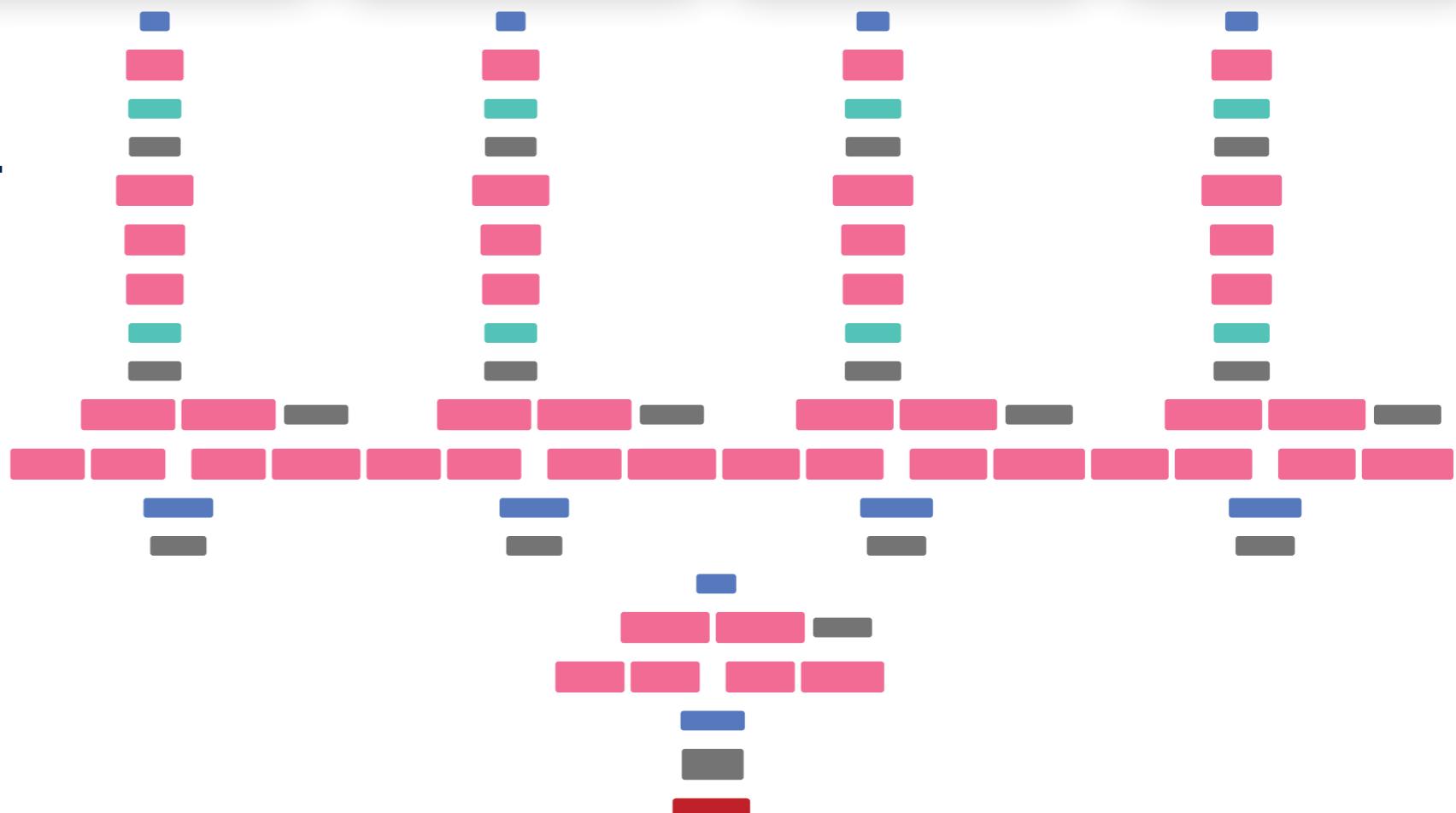


Using the existing reconstruction.

Improved CNN network to optimize
for running time

Modified to take 4 views (event +
prong)

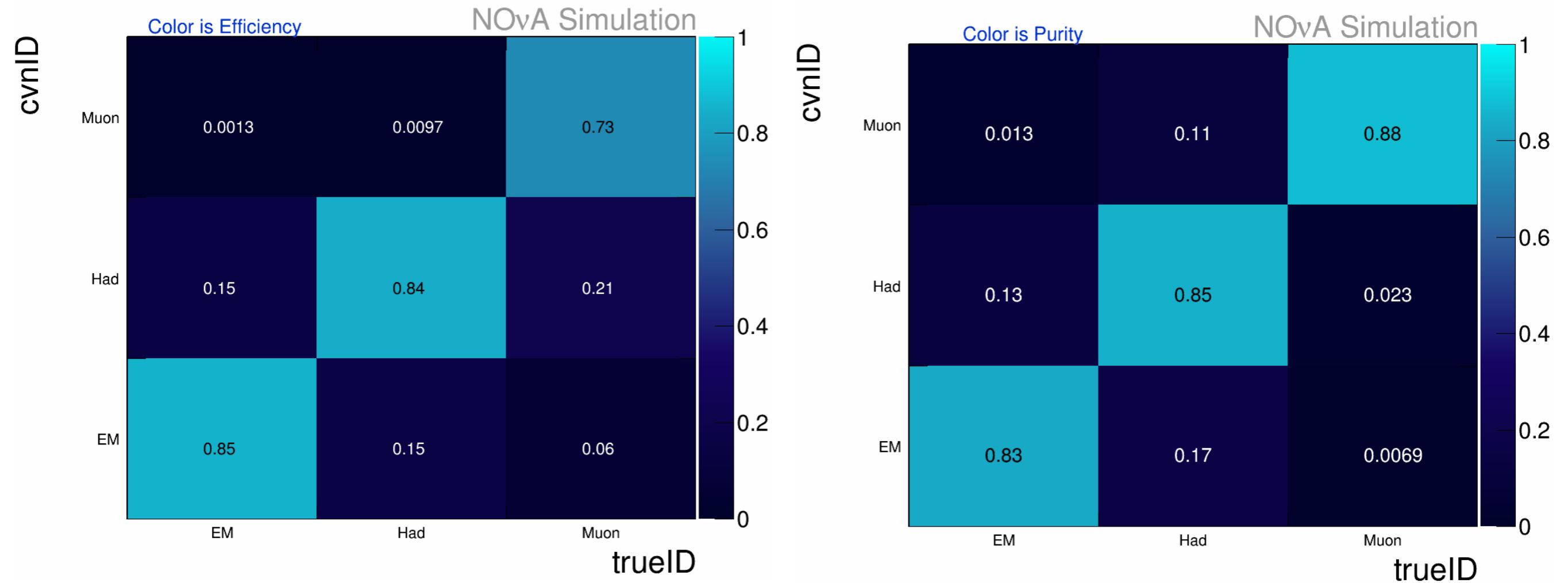
Trained on clusters from all events
above some minimum purity.



Cluster Classification Matrix

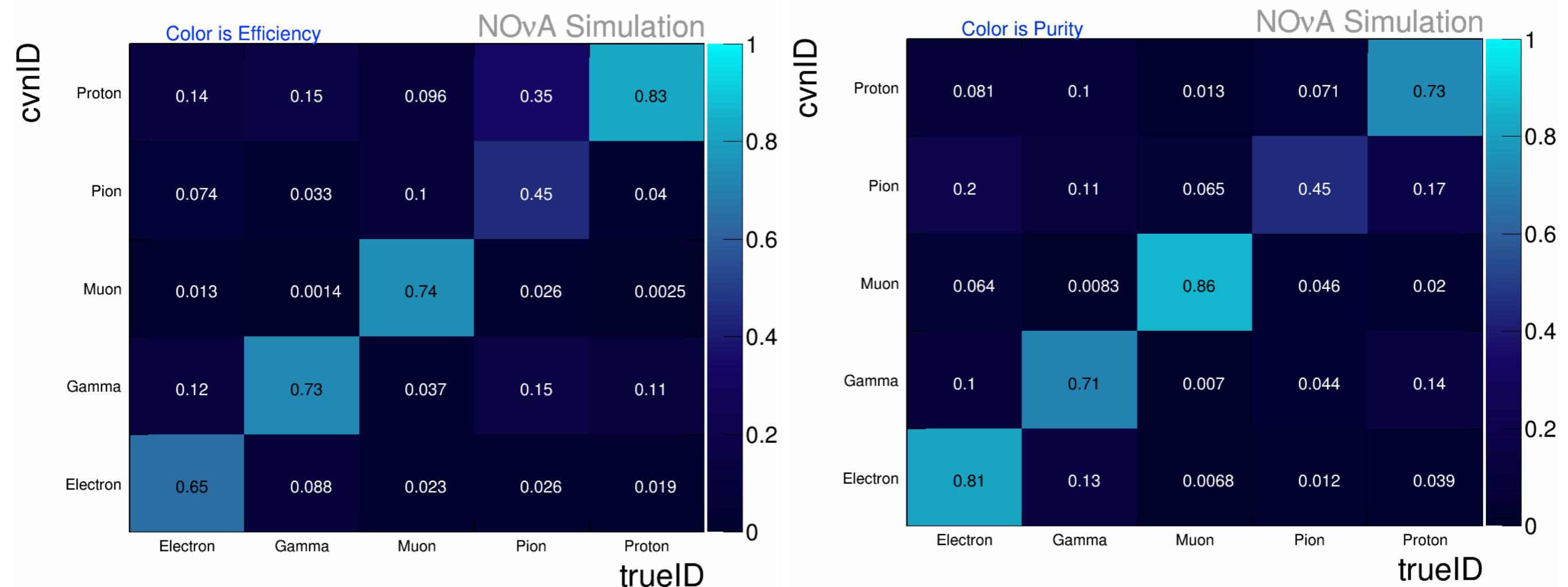
Broad categories separate electromagnetic and hadronic contributions as well as muons.

This network classifies clusters below some length threshold.



Cluster Classification Matrix

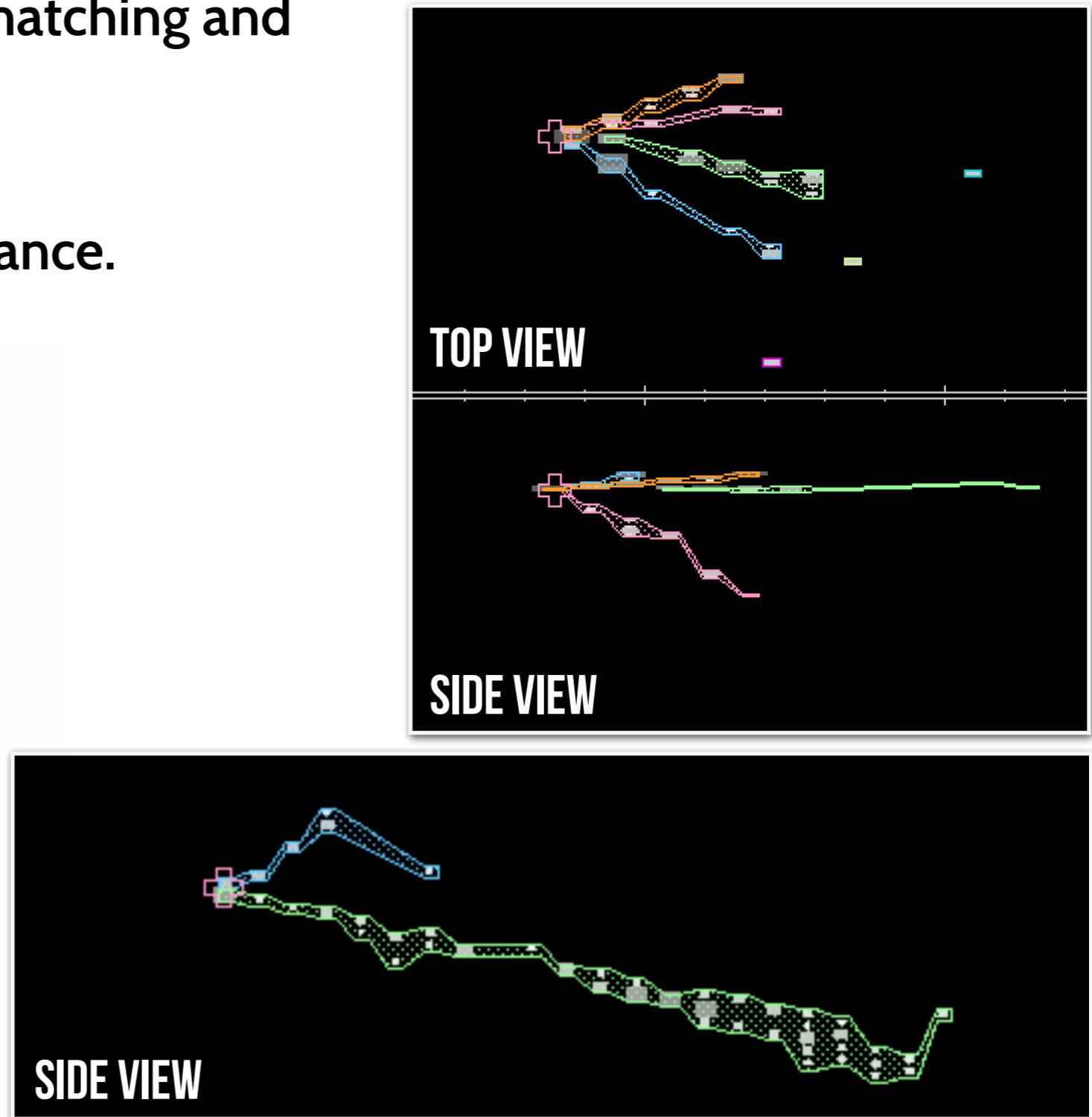
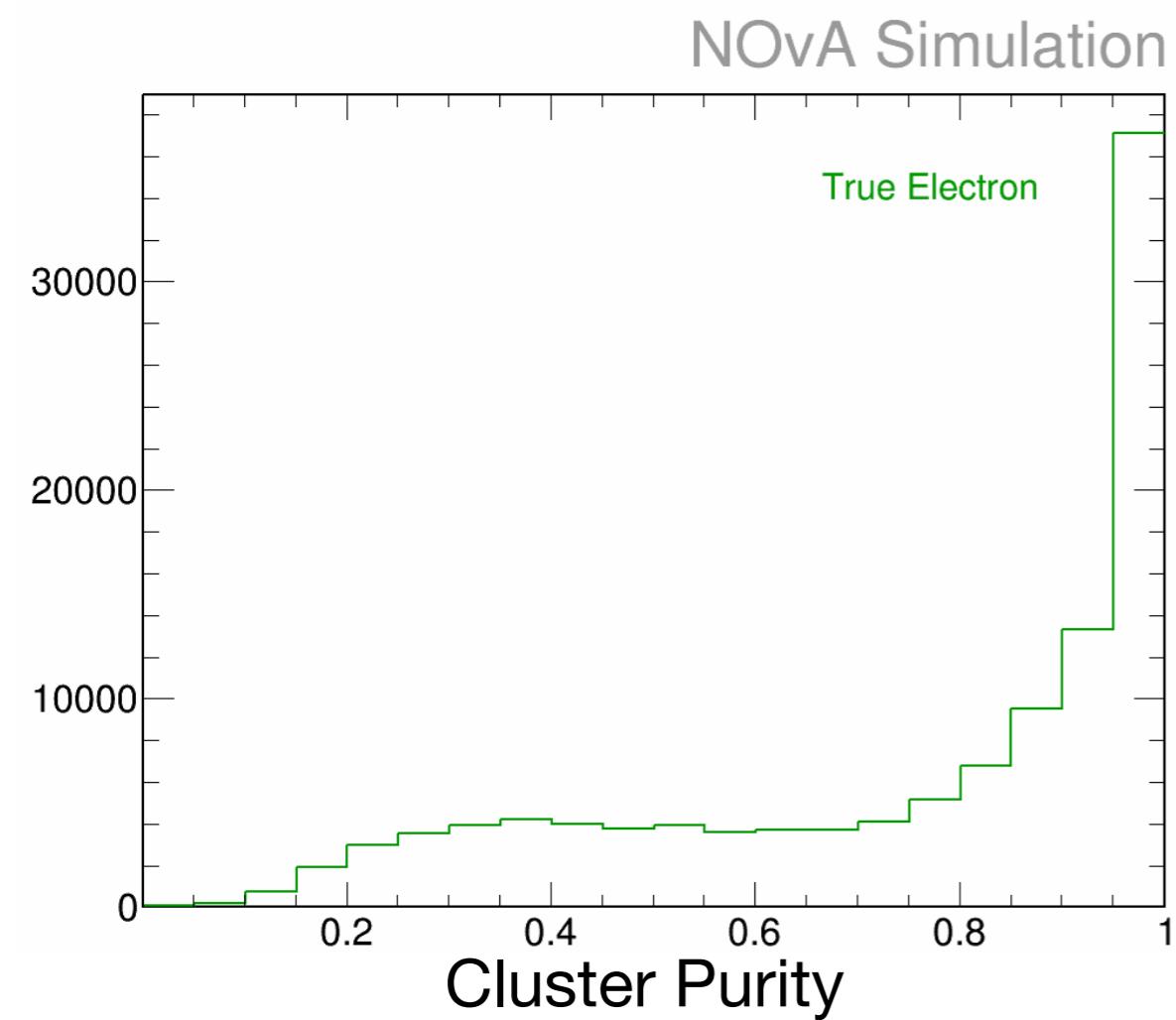
Full set of training categories for single particle tagging.



Caveats from the reconstruction

Prong quality depends on view matching and vertex reconstruction.

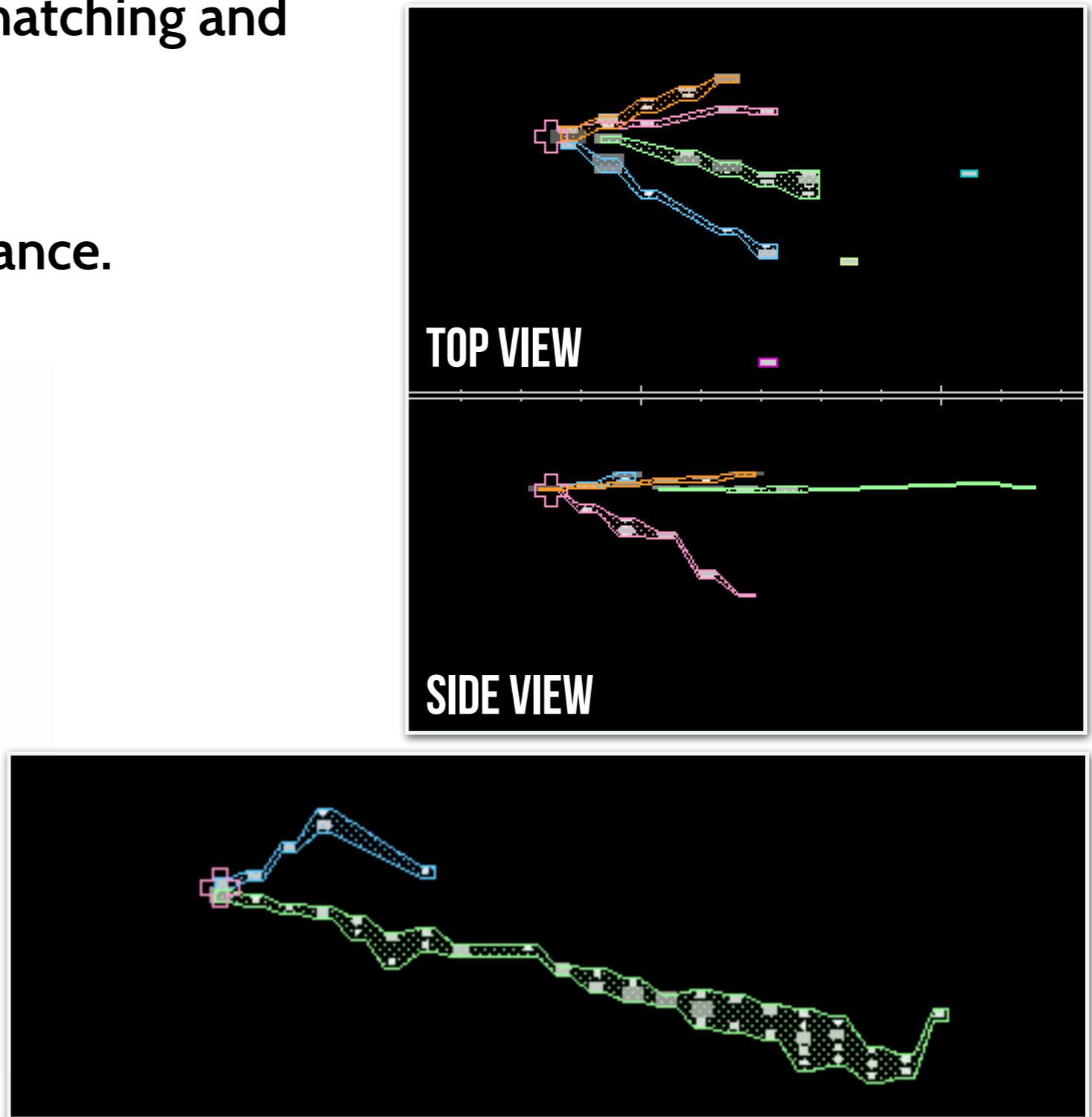
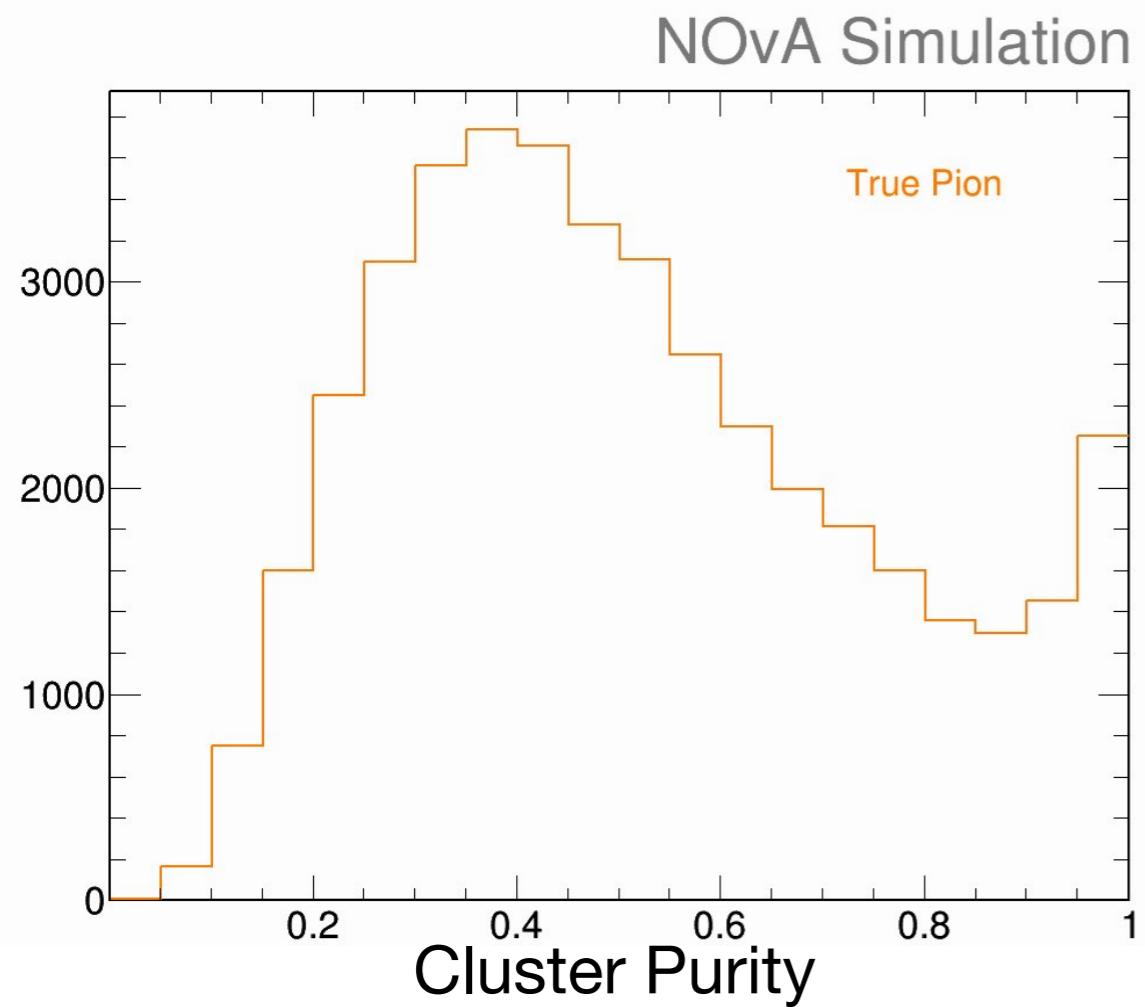
Purity impacts classifier performance.



Caveats from the reconstruction

Prong quality depends on view matching and vertex reconstruction.

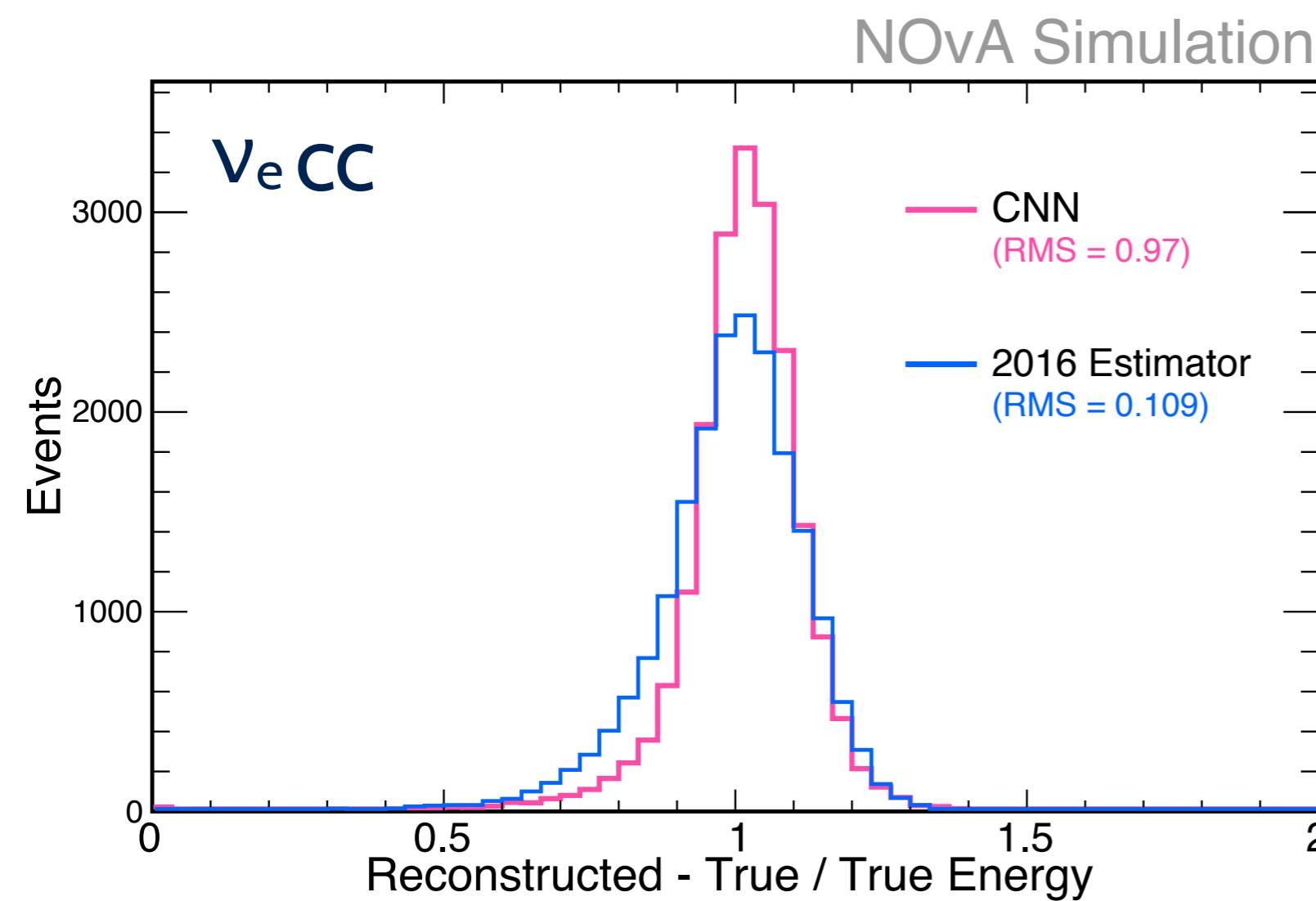
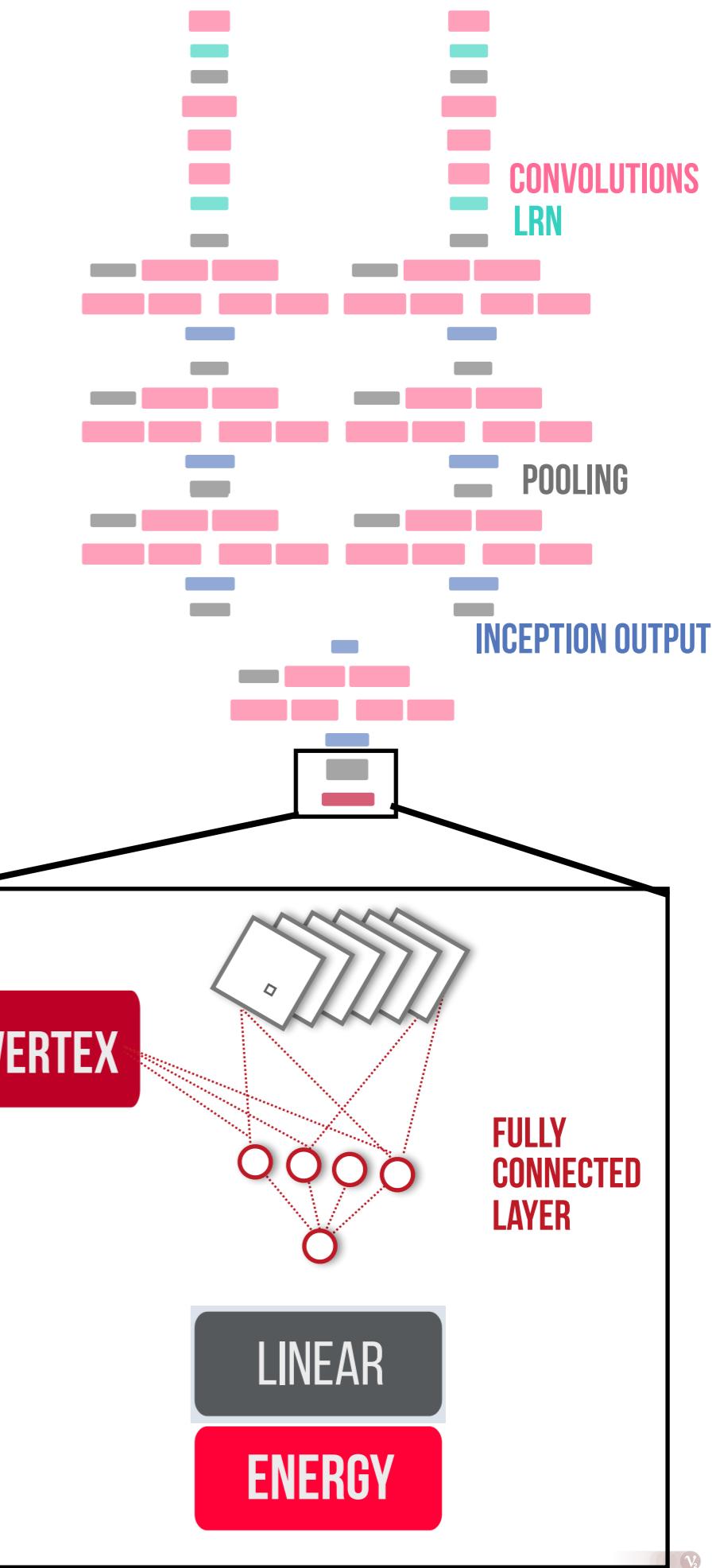
Purity impacts classifier performance.



CNNs for Energy Reconstruction

The target is Energy instead of PID values

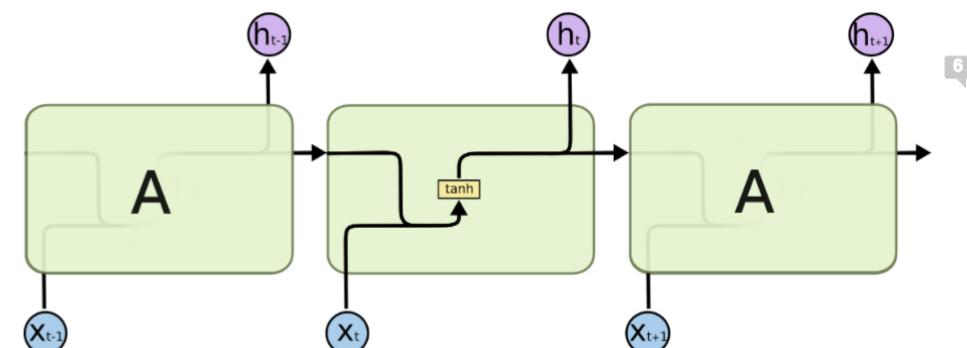
Incorporate information from the reconstruction at the fully connected stage.



LSTMs for Energy Estimation

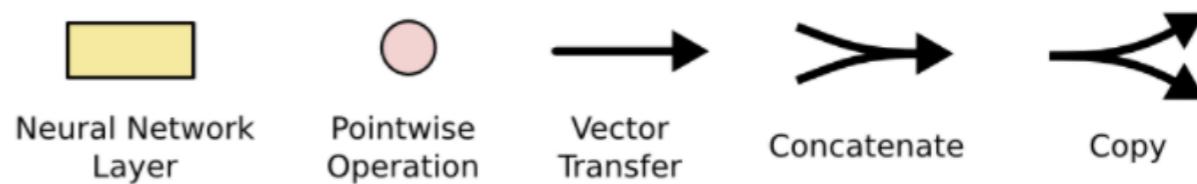
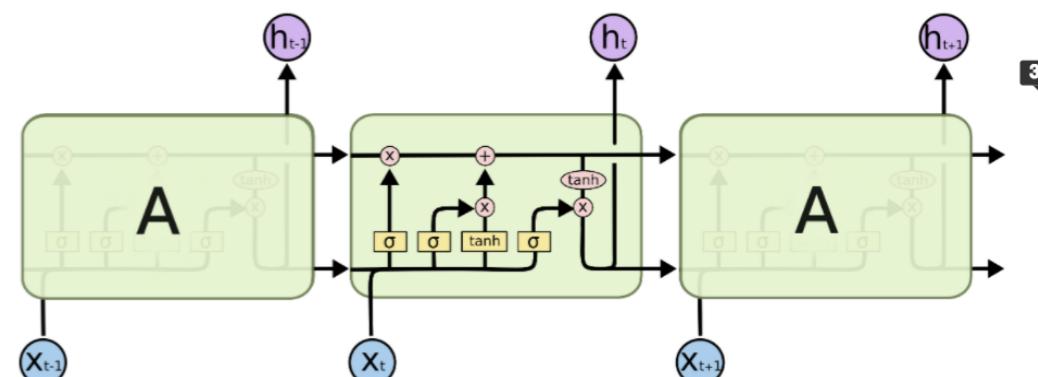
Recurrent Neural Networks:

Sequential network using the current state of the system + the output from last iteration.



LSTMs are RNN + Cell State

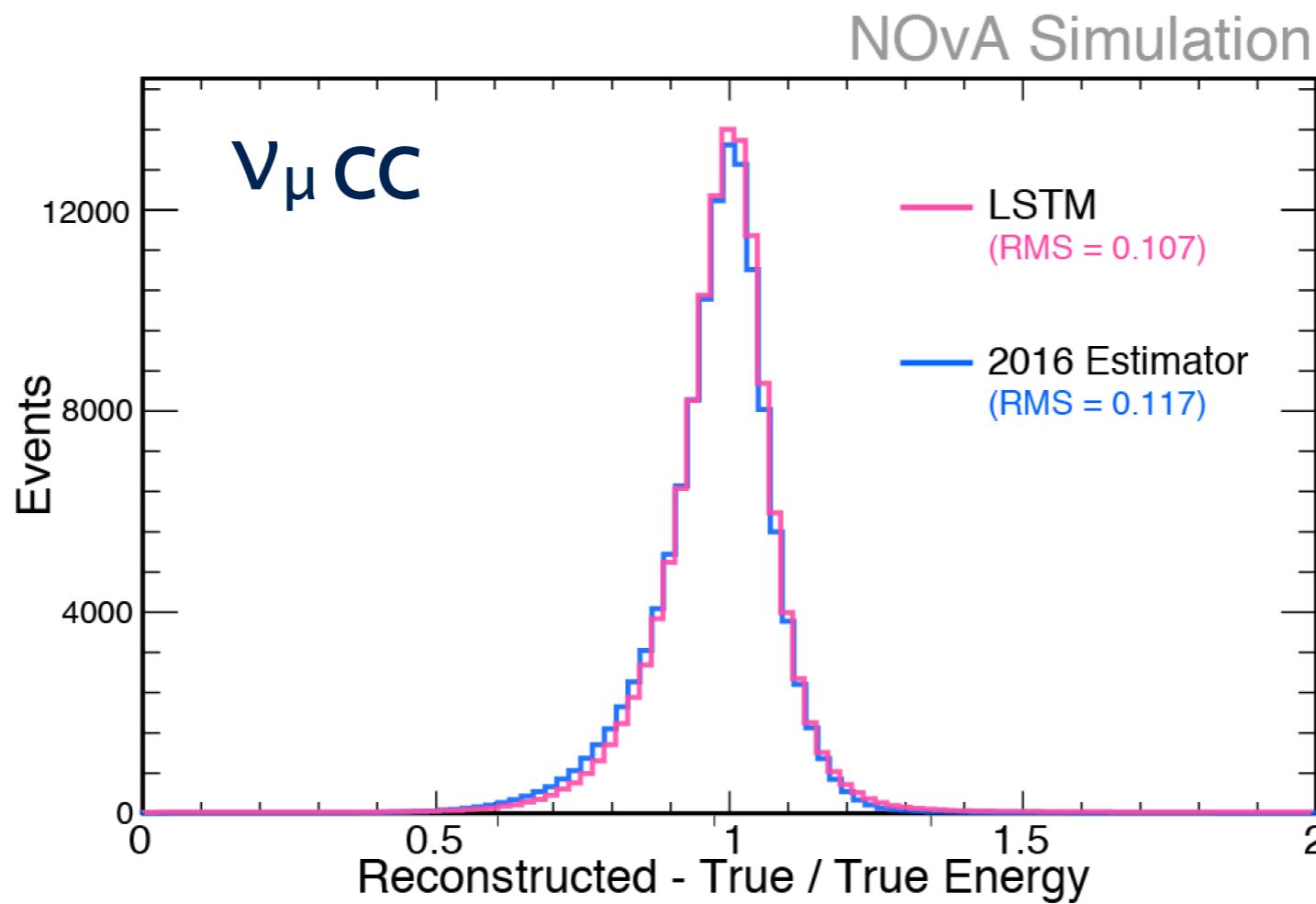
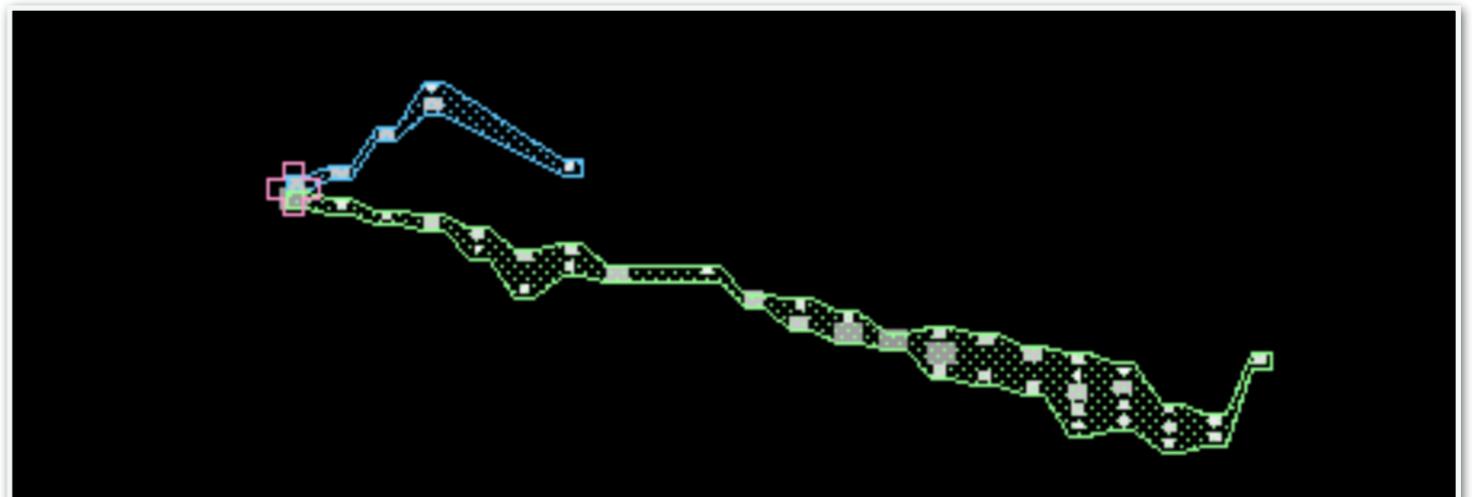
Long Term - Short Term Memory cell states are secondary paths which keep long term memory of the system.



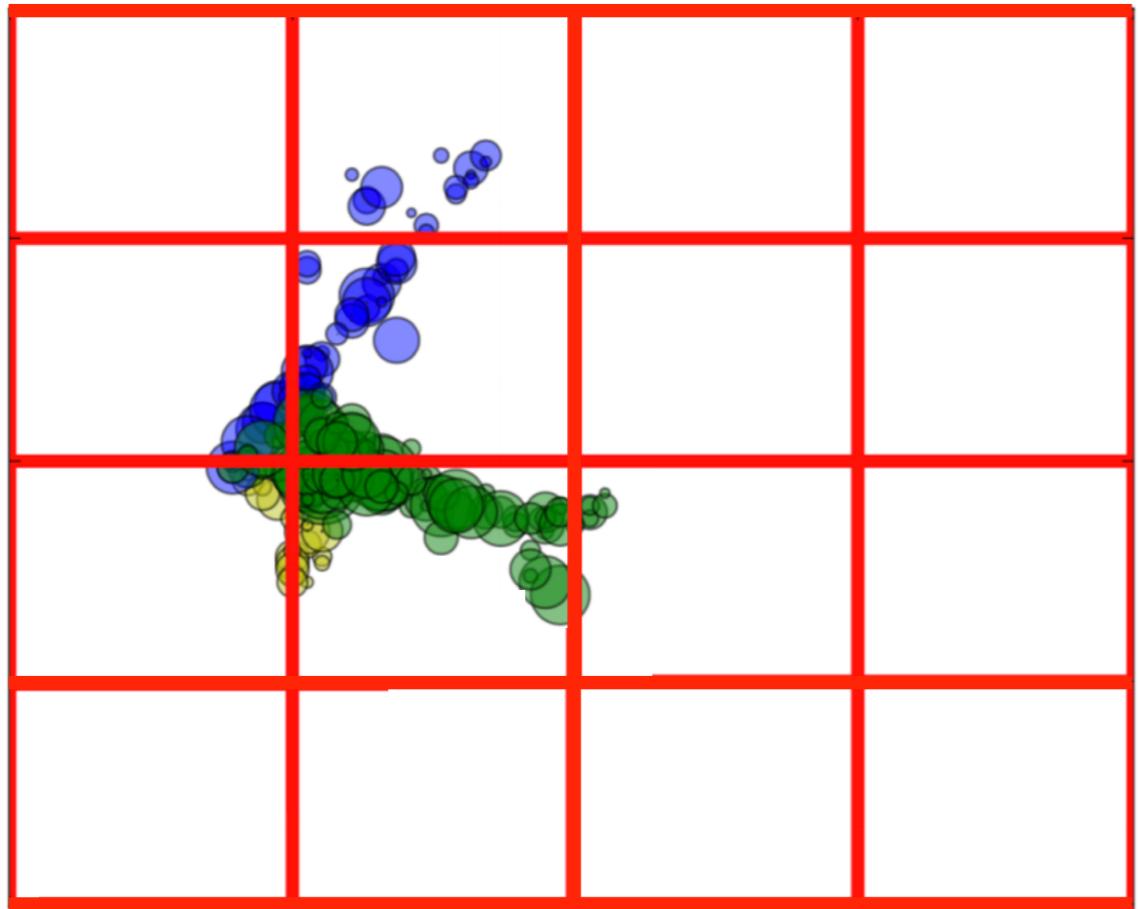
OnGoing RNNs for Energy Estimation

Input for each cluster:

- Direction and length
- Cluster Particle ID
- Energy estimates
- Event calibrated energy

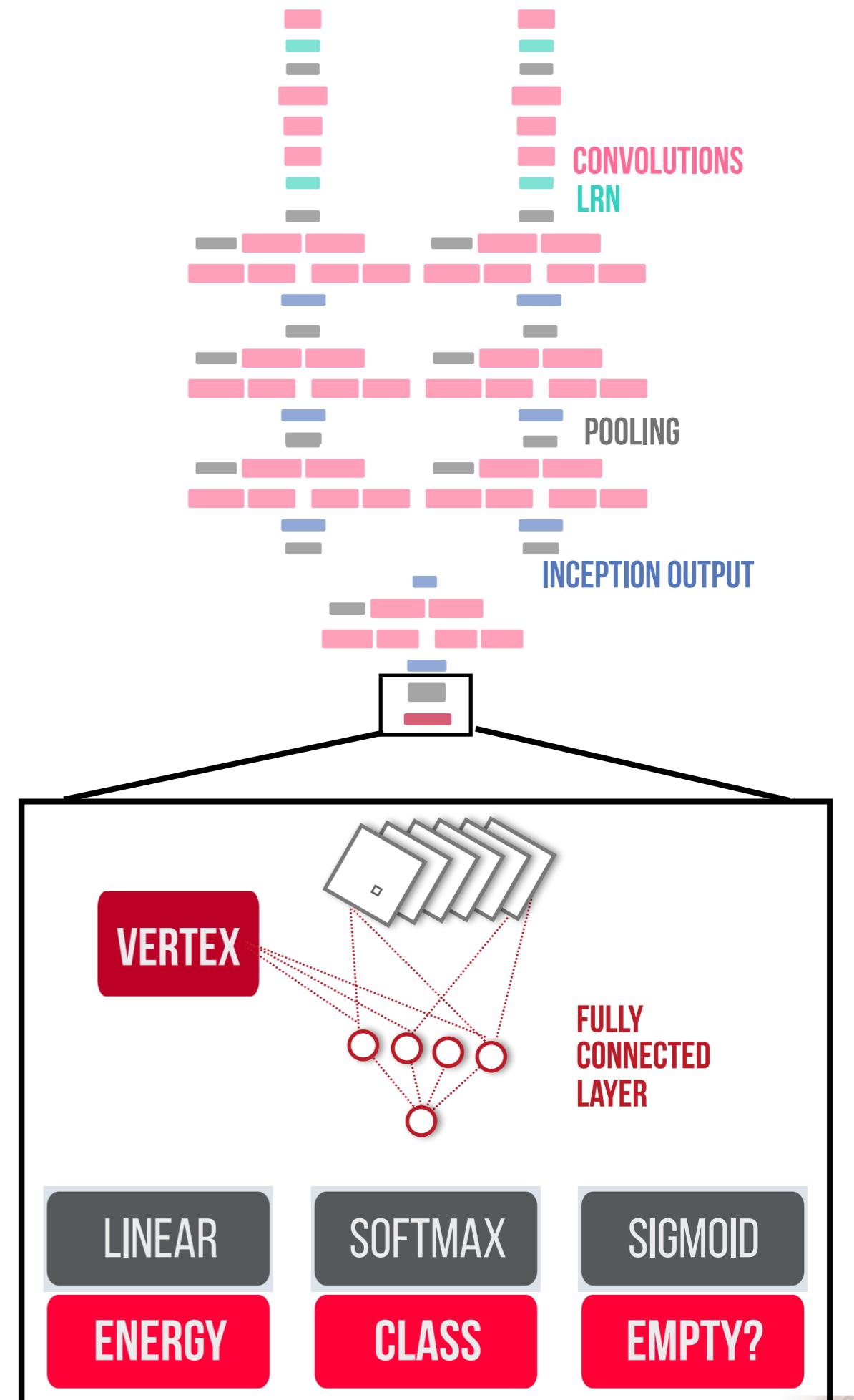


Multi-target DNNs

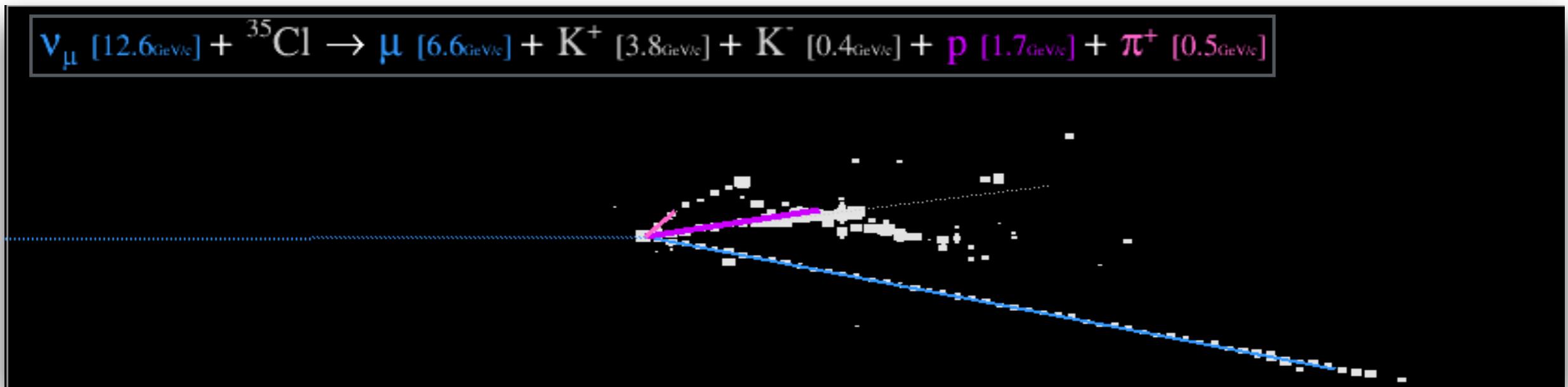


We are exploring CNN architectures which allow for multiple purposes.

Preliminary tests give a competitive energy resolution.



Summary of Ongoing Work



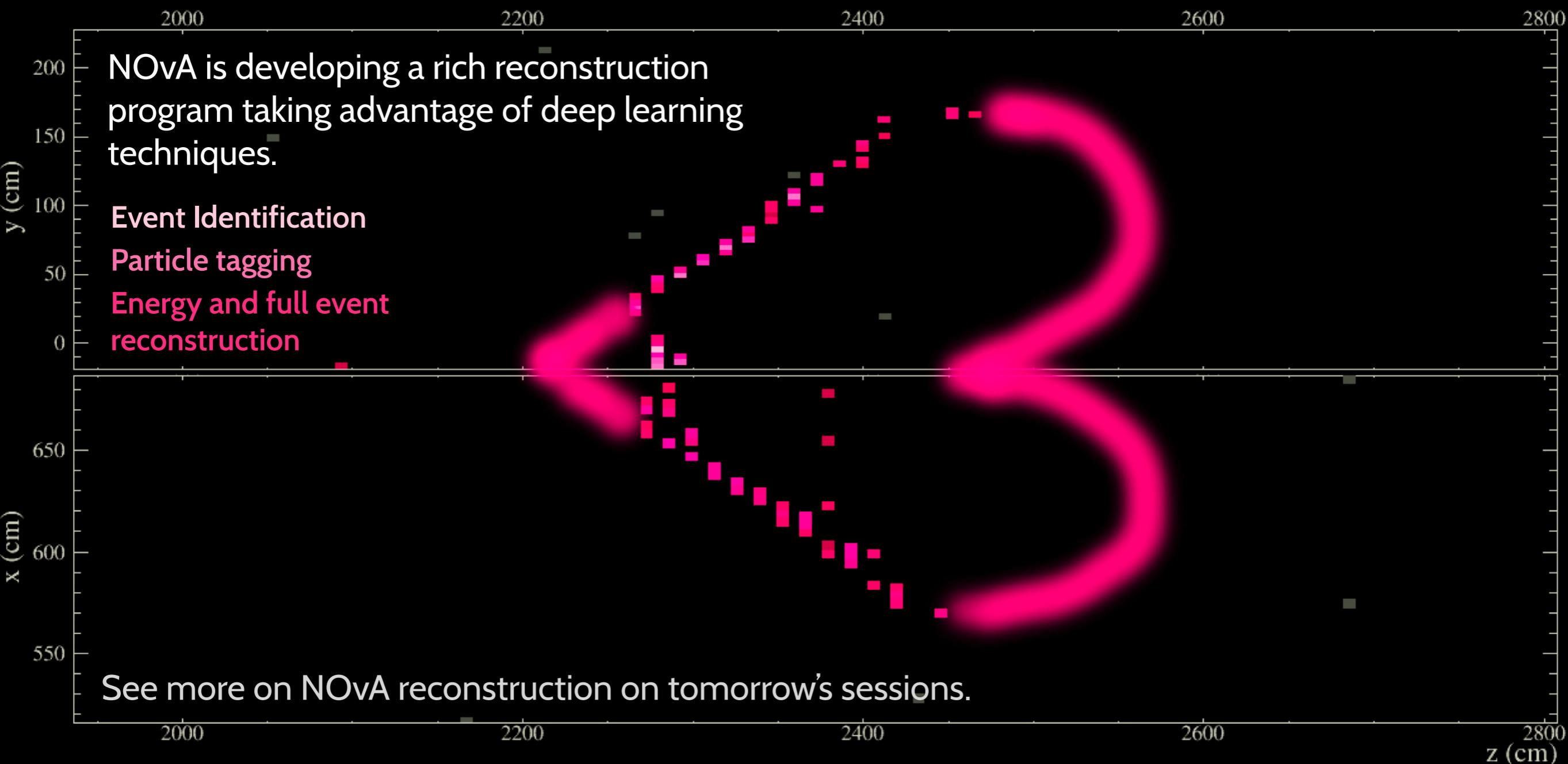
Particle by particle tagging is currently in use for energy estimation and selection in cross sections analyses.

Energy reconstruction is under development from multiple, complimentary approaches.

More Implementations like **vertex reconstruction** and full event **identification + clustering** (hit by hit) are also being developed



Summary



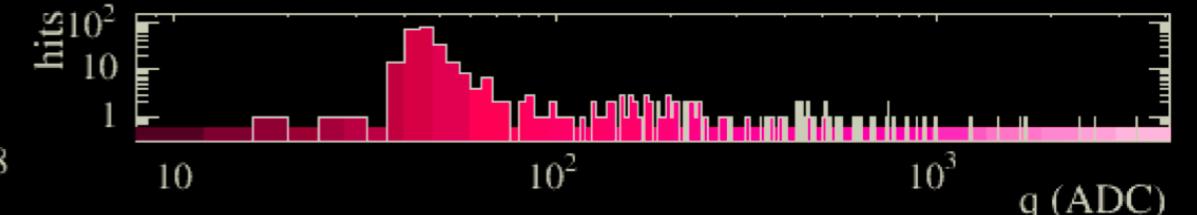
NOvA - FNAL E929

Run: 15330 / 4

Event: 11978 / --

UTC Fri May 23, 2014

17:30:2.632293184



Backup

These are not the slides you're looking for

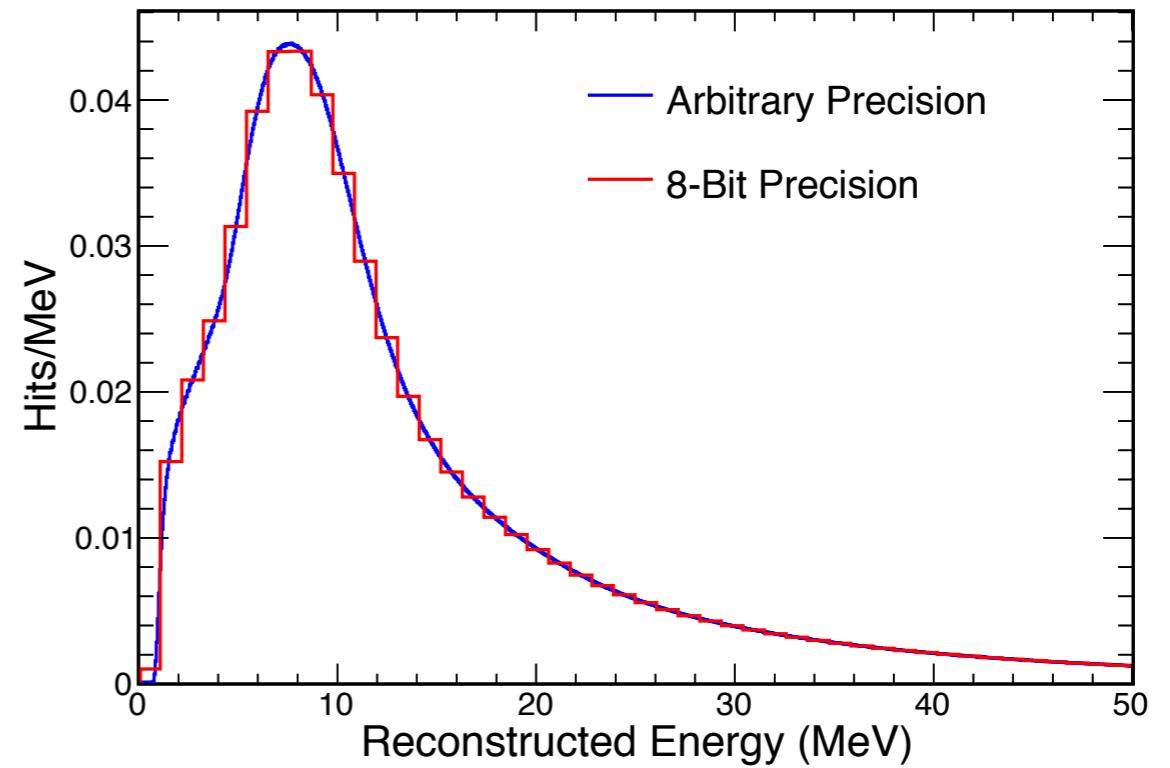
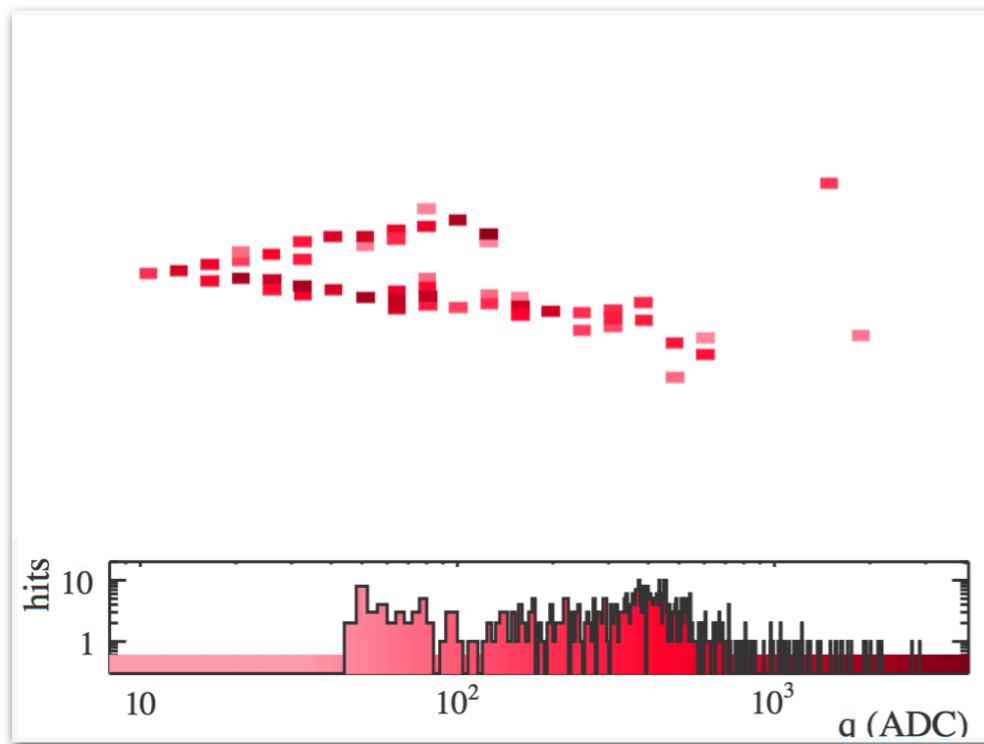


Event ID with Convolutional Neural Networks

Premise: Let a deep learning network extract features and draw correlations.

Disentangle the identification from reconstruction.

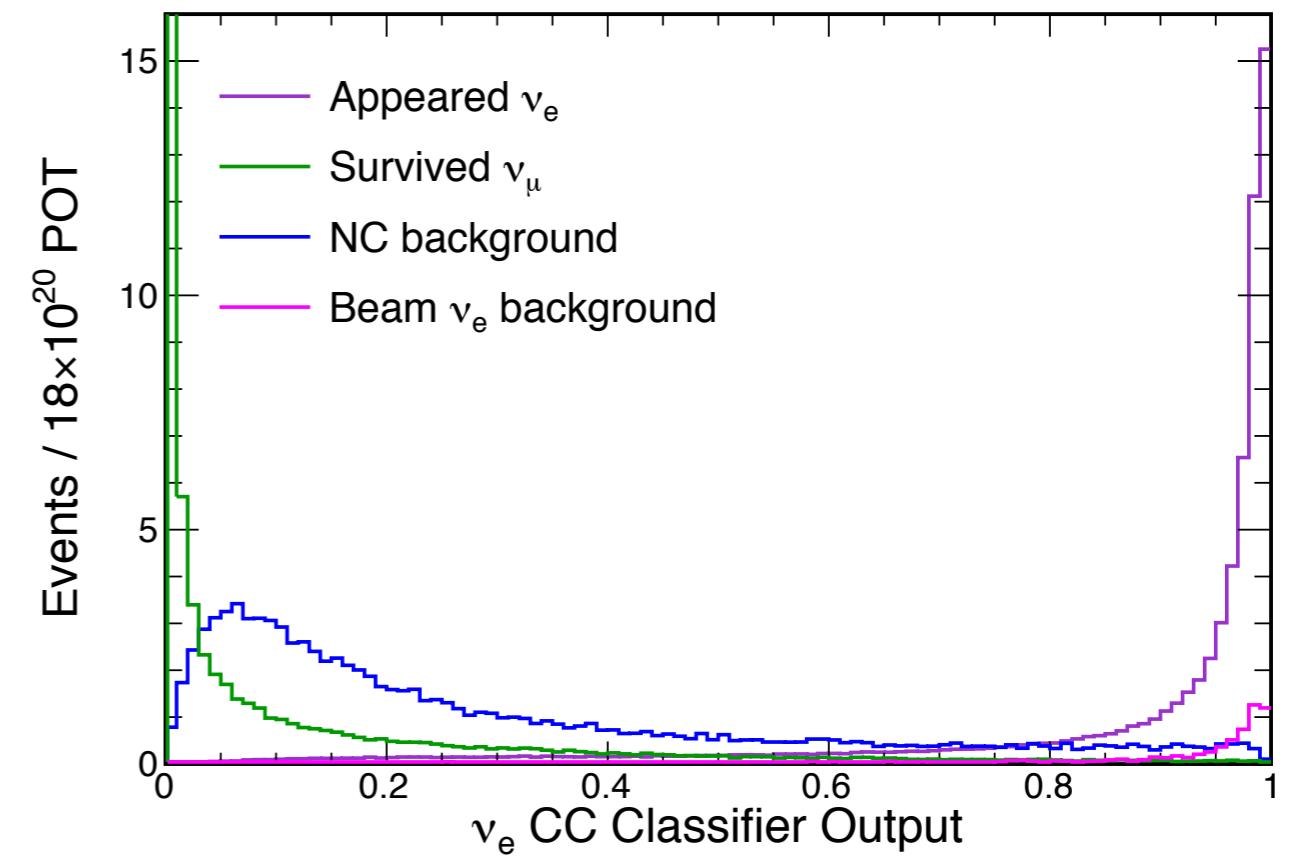
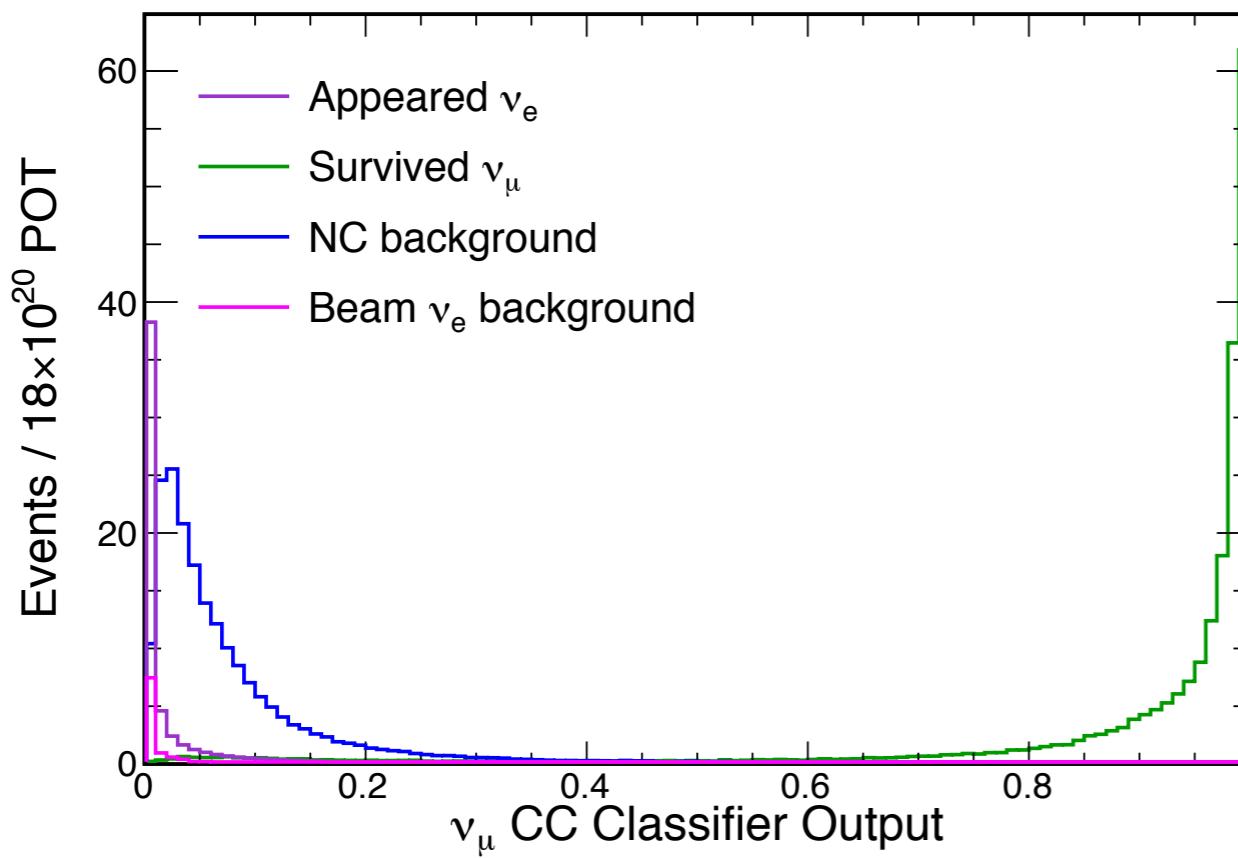
In practice: Use “images” of events to train a CNNs to identify neutrino flavor.



Classifier output

One PID value for each category, normalized to sum 1 over all possible labels.

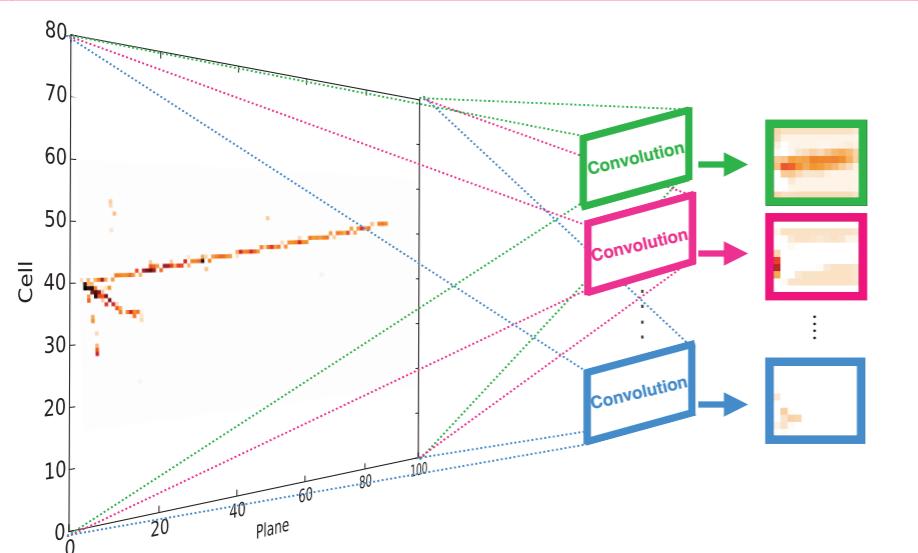
In principle, this means one can extract more information than a single PID value.



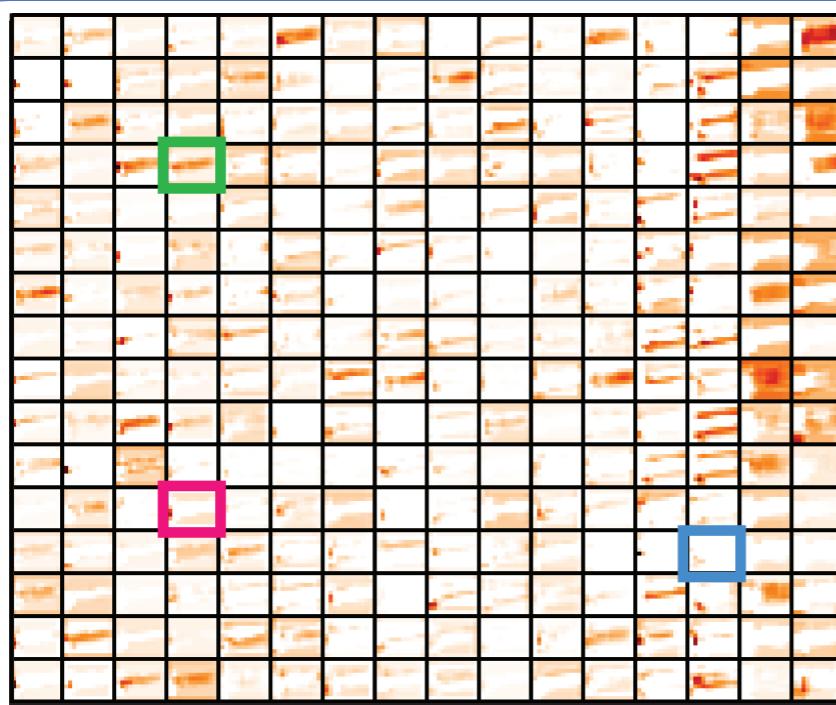
CVN Network Components

Convolutional Visual Network

CONVOLUTION



INCEPTION OUTPUT



CONVOLUTIONS

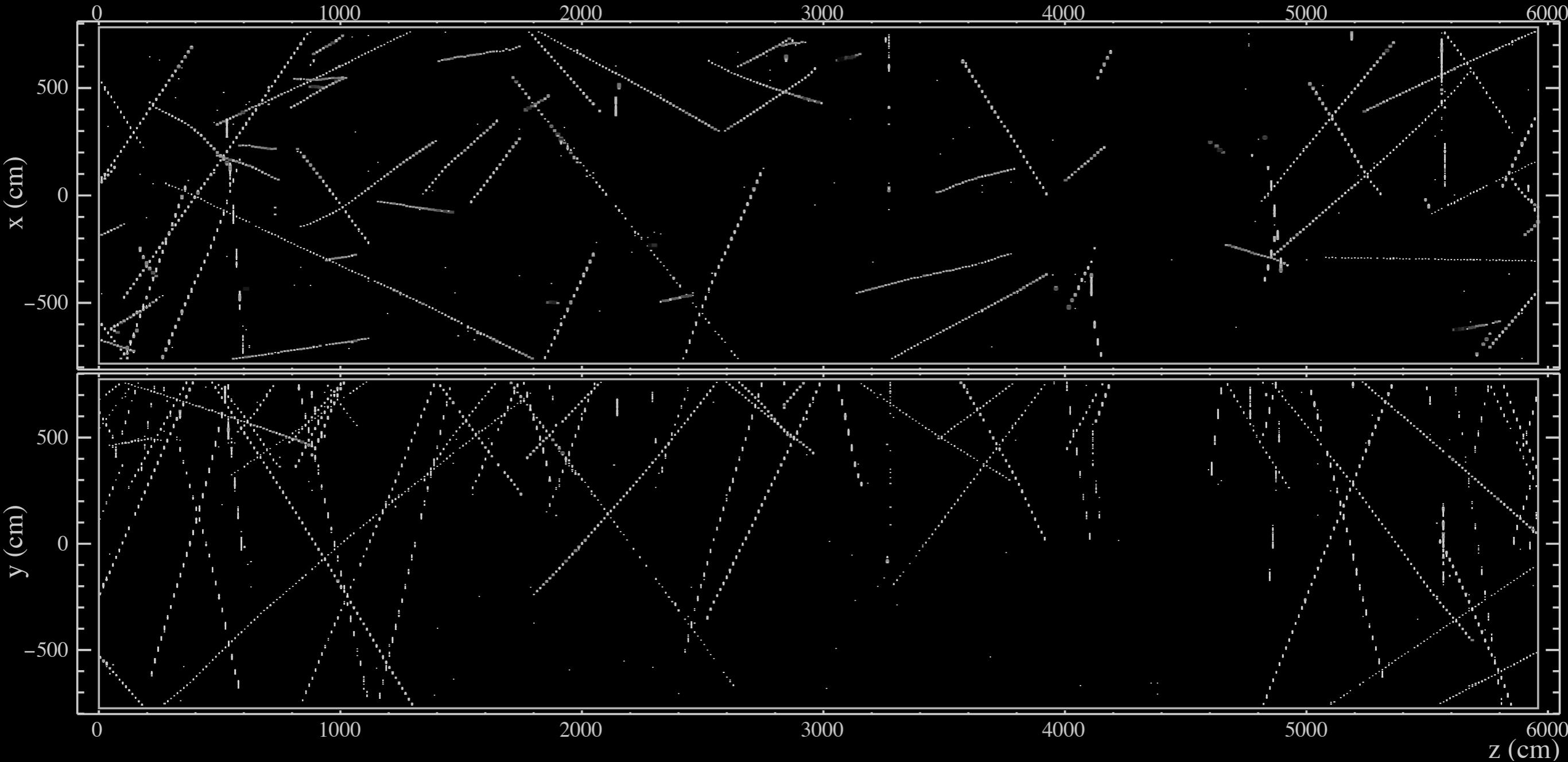
LRN

POOLING

INCEPTION OUTPUT

Isolating neutrino interactions

The first step in our reconstruction is dividing an **event** (550 μs of data)



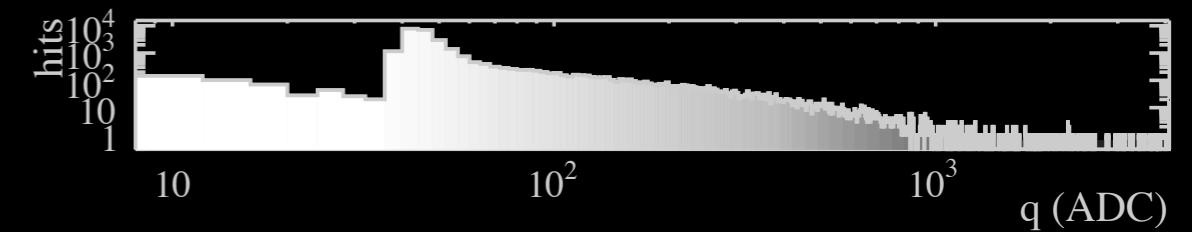
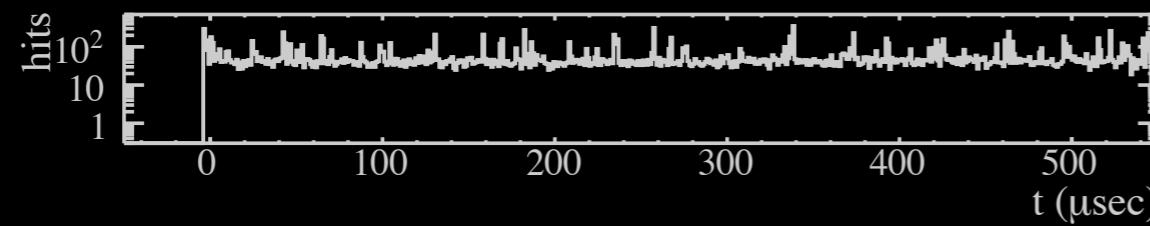
NOvA - FNAL E929

Run: 19193 / 13

Event: 188331 / --

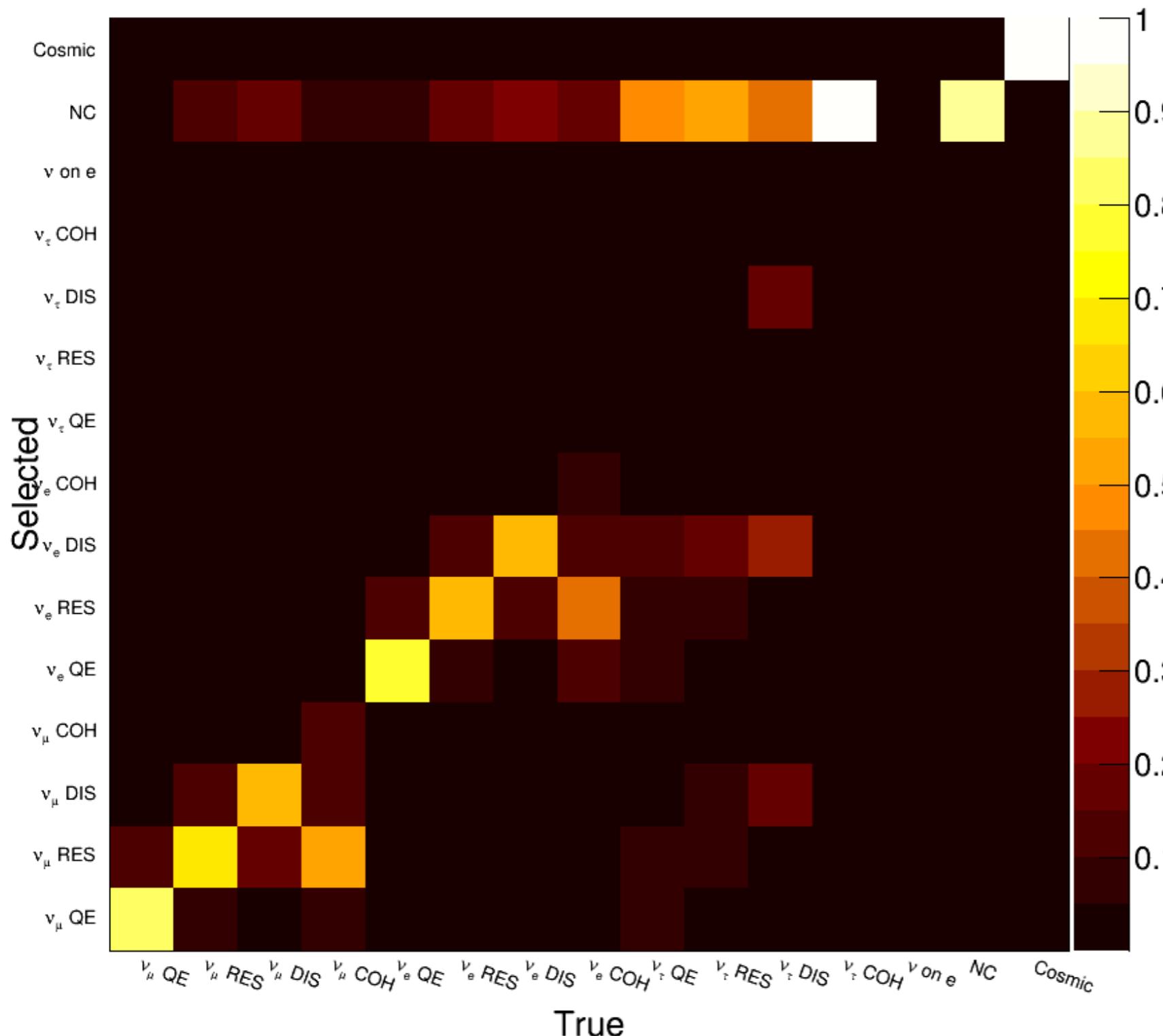
UTC Fri Mar 27, 2015

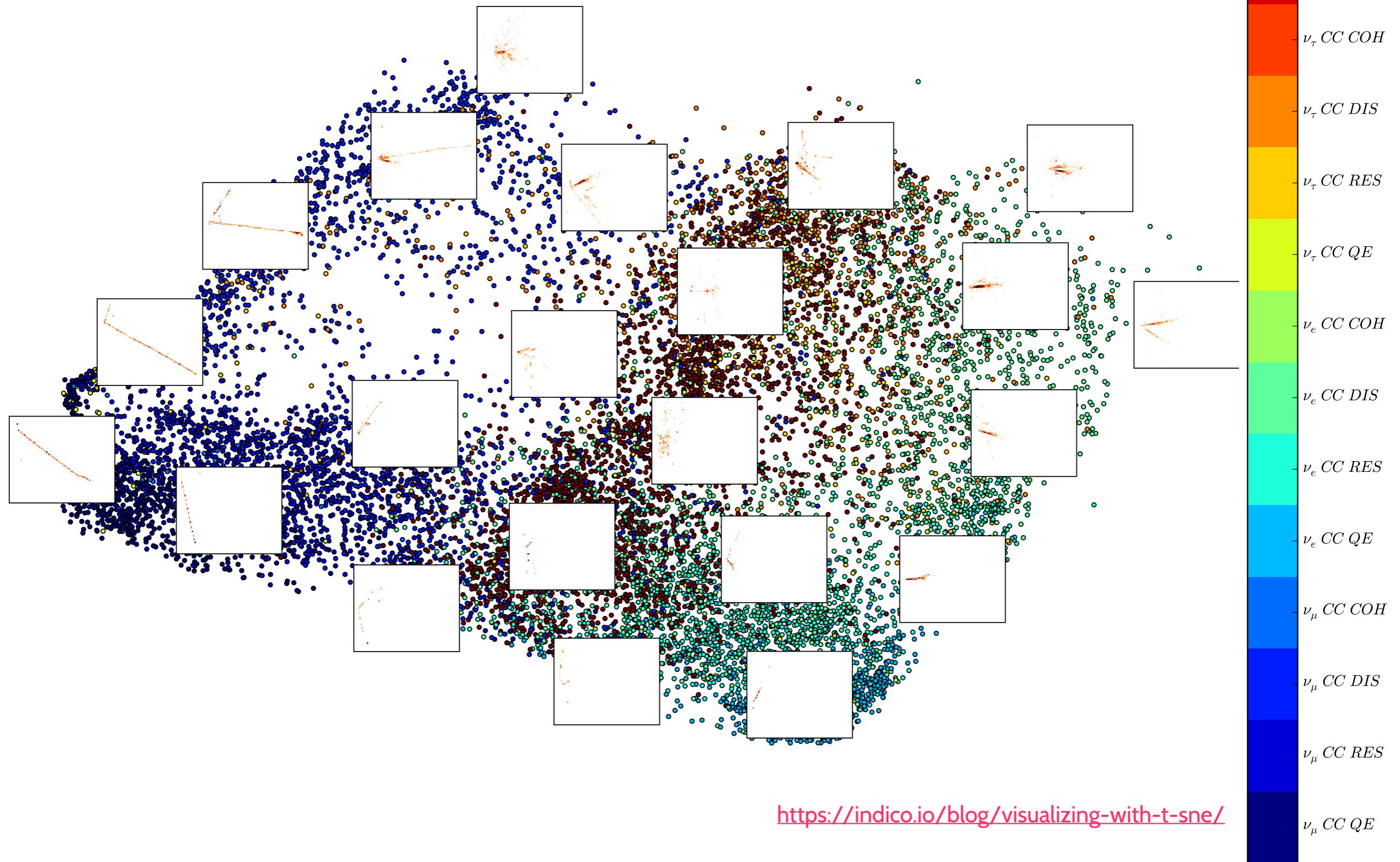
09:44:53.281953920



Correlation Matrix

NOvA Simulation

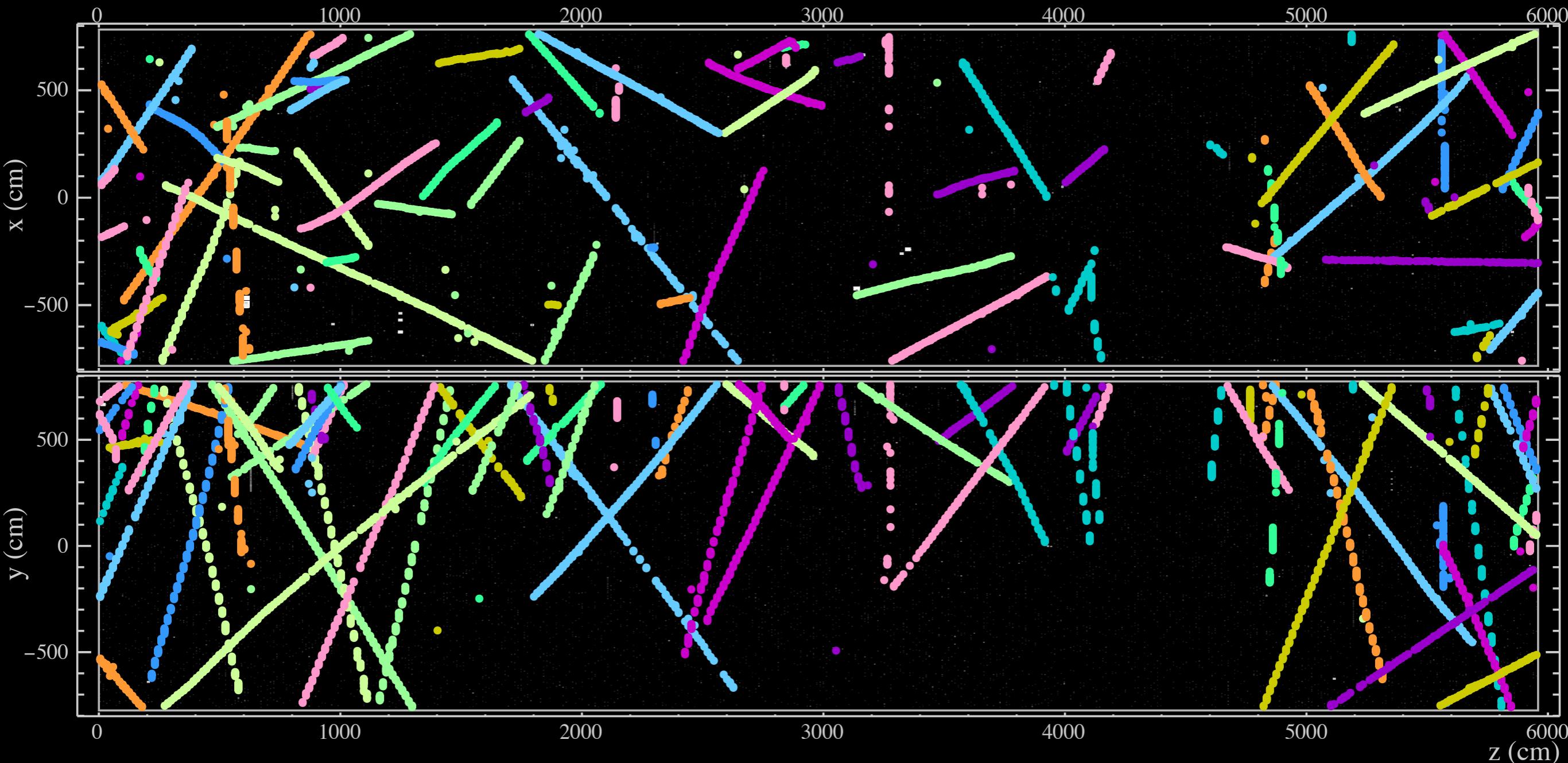




<https://indico.io/blog/visualizing-with-t-sne/>

Isolating neutrino interactions

The first step in our reconstruction is dividing an **event** ($550 \mu\text{s}$ of data) into slices (groups of hits with some time and space coincidence)



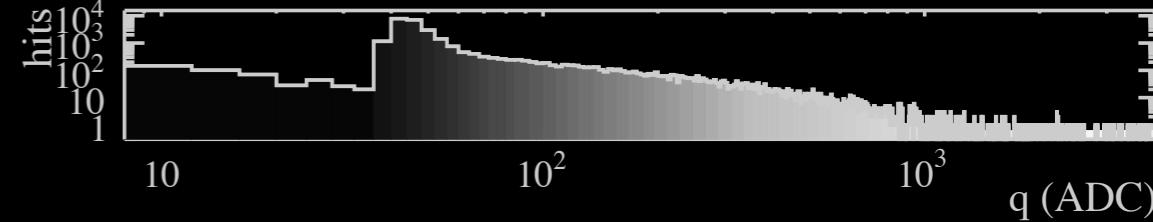
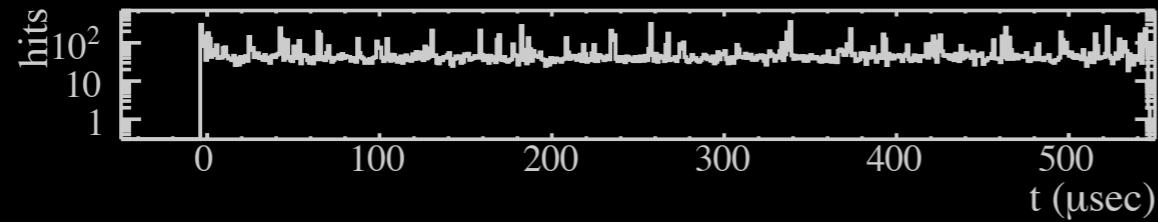
NOvA - FNAL E929

Run: 19193 / 13

Event: 188331 / --

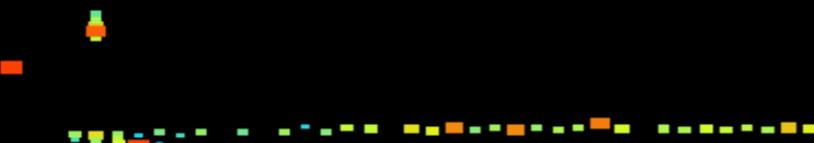
UTC Fri Mar 27, 2015

09:44:53.281953920

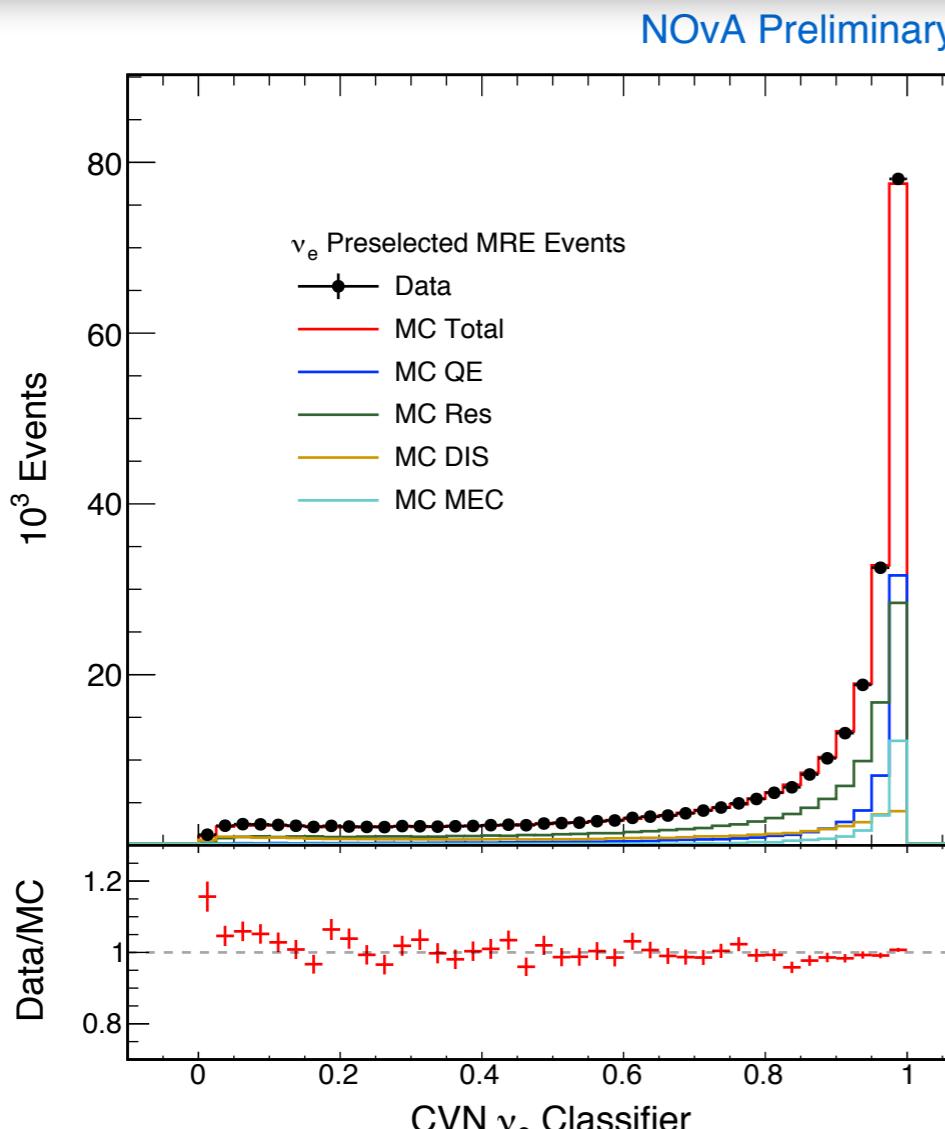


CVN Performance On Real Data

DATA



DATA $\mu \longleftrightarrow e$



MRE (Muon Removed - Electron):

Select a muon neutrino interaction with traditional ID methods.

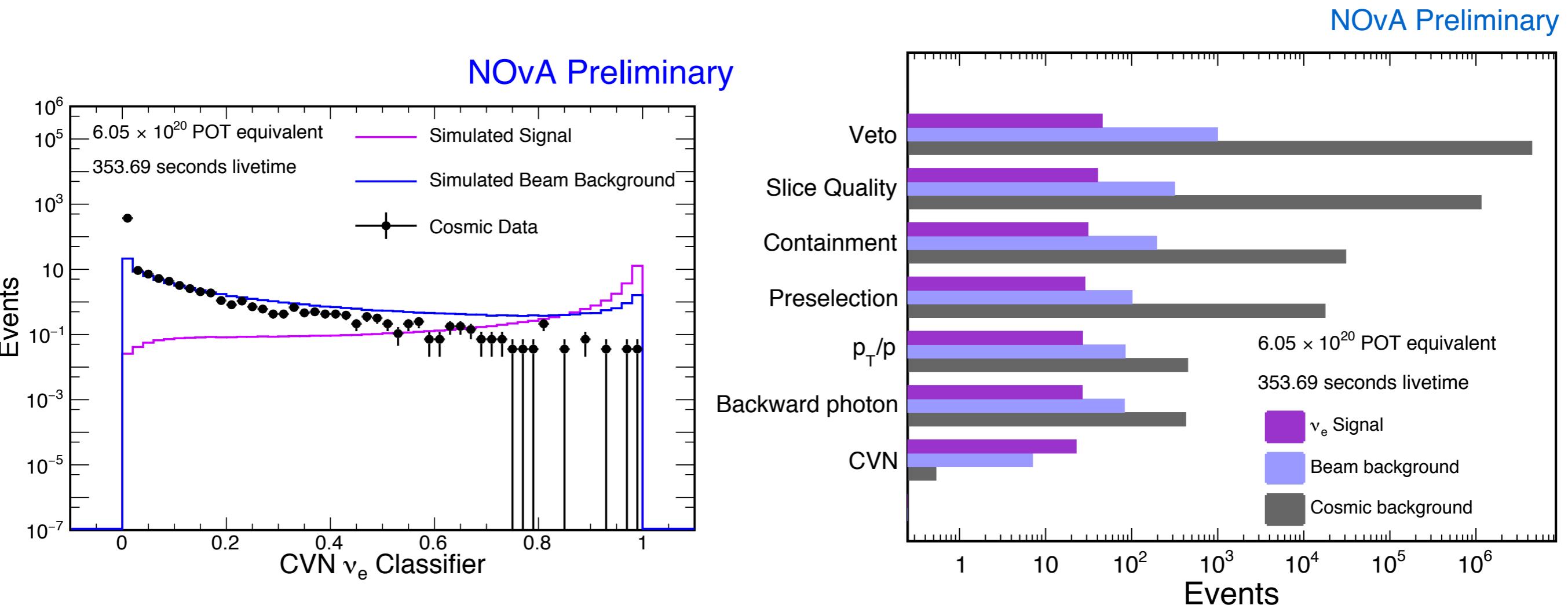
Remove the muon hits and replace them with a single simulated electron of matching momentum.

Data/MC comparisons show less than 1% difference in efficiency.

PID	Sample	Preselection	PID	Efficiency	Efficiency diff %
CVN	Data	262884	188809	0.718222	-
	MC	277320	199895	0.720809	-0.36%

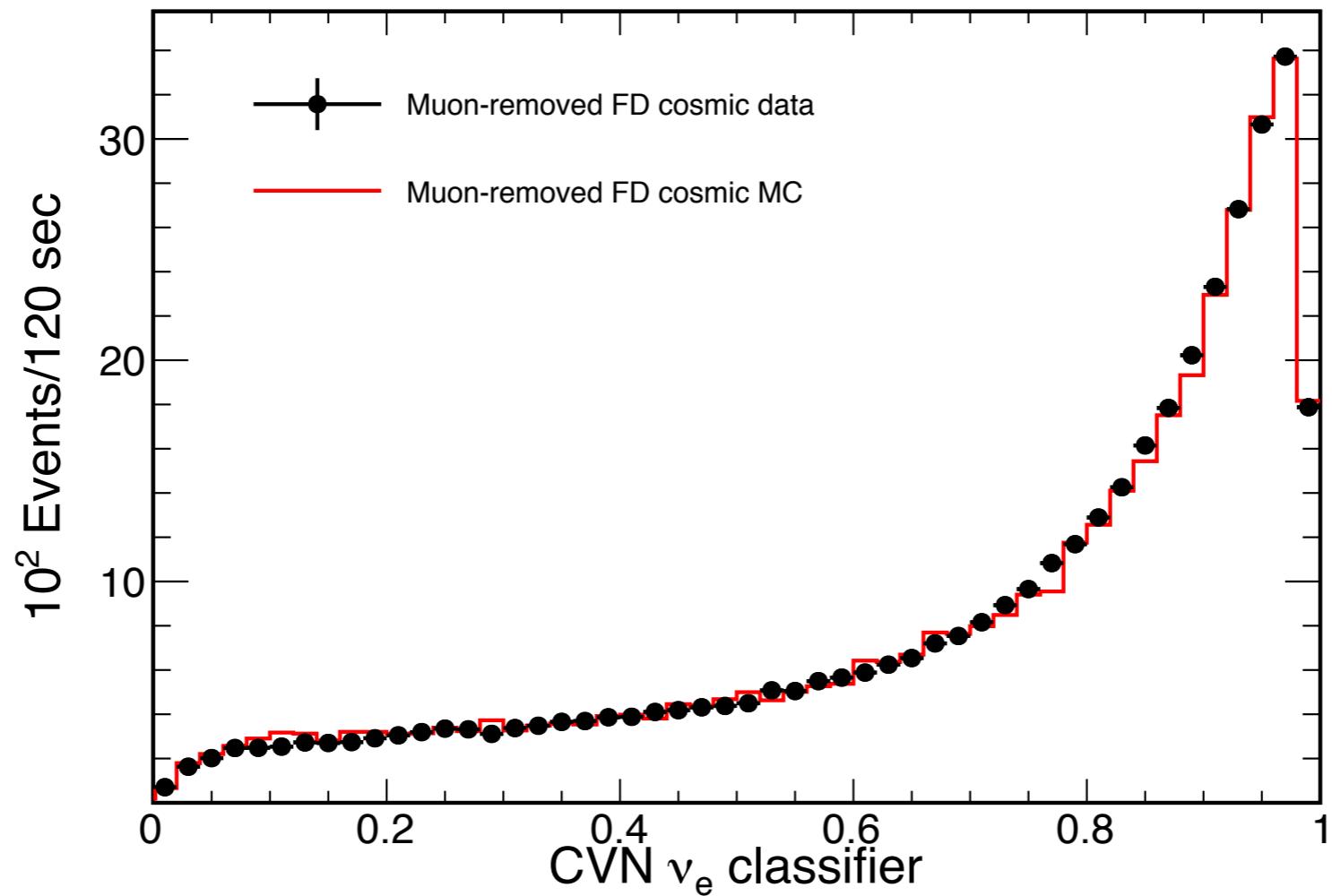


Performance on Cosmic Background



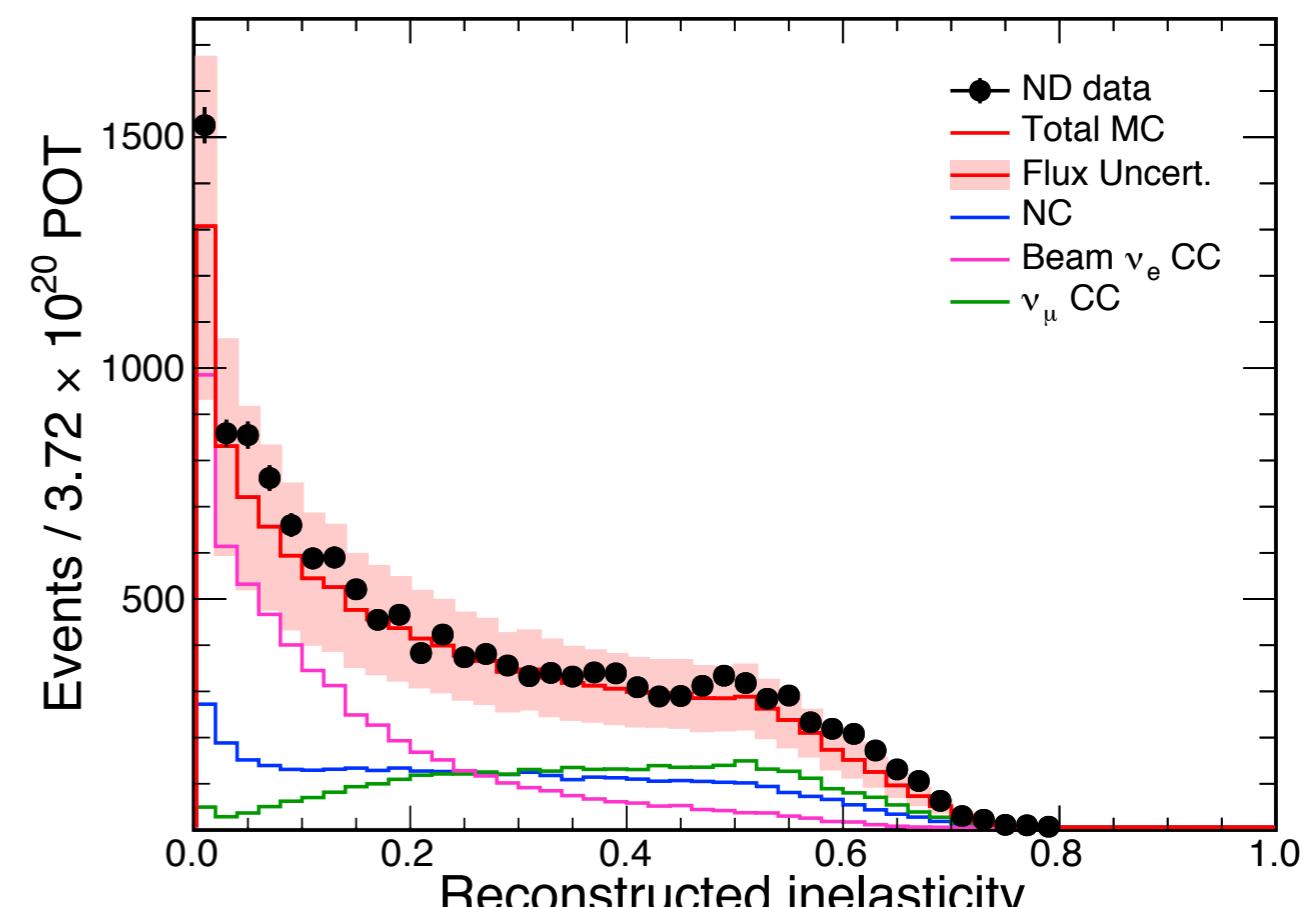
Data Driven Tests - MRBrem

NOvA Preliminary

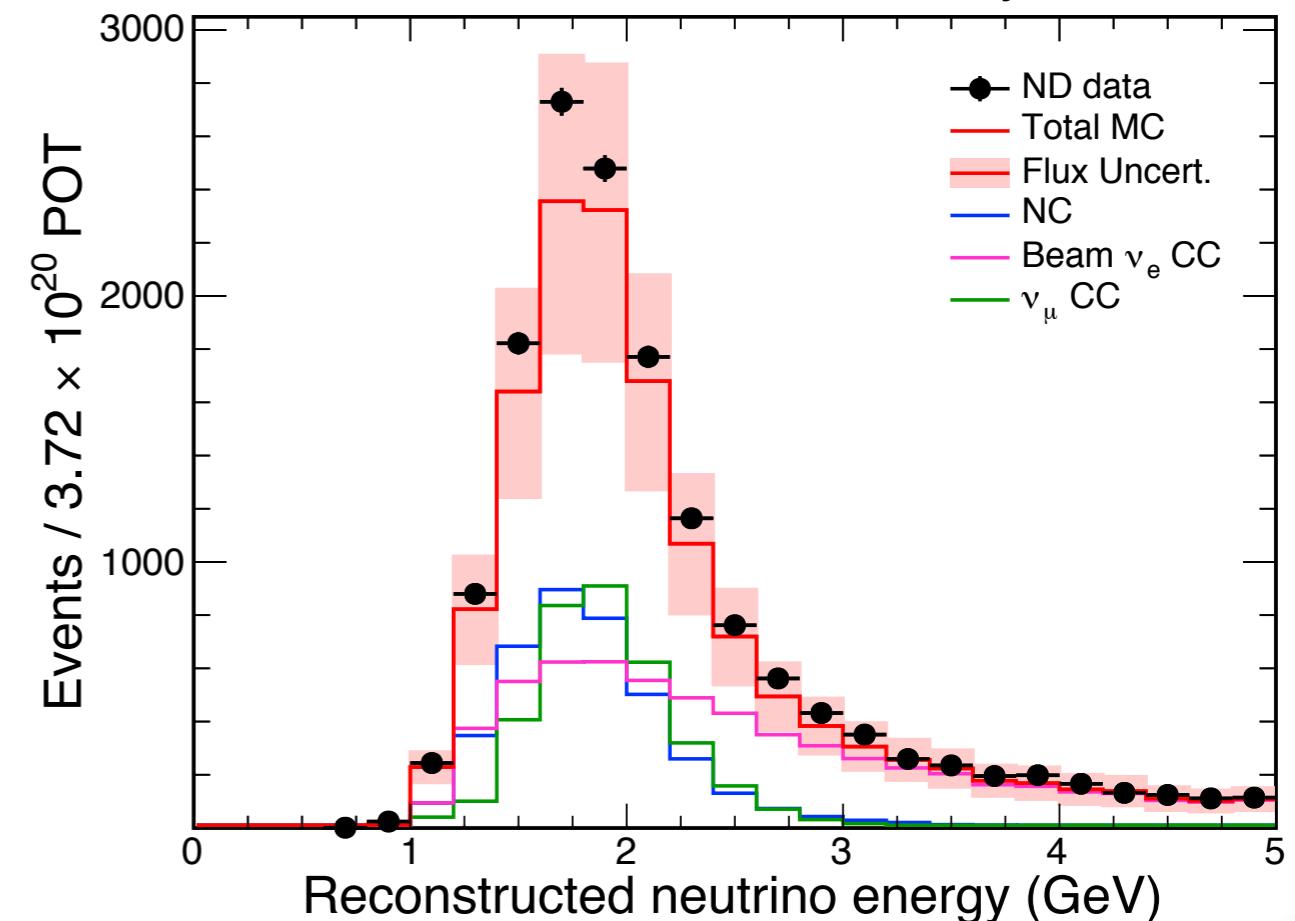
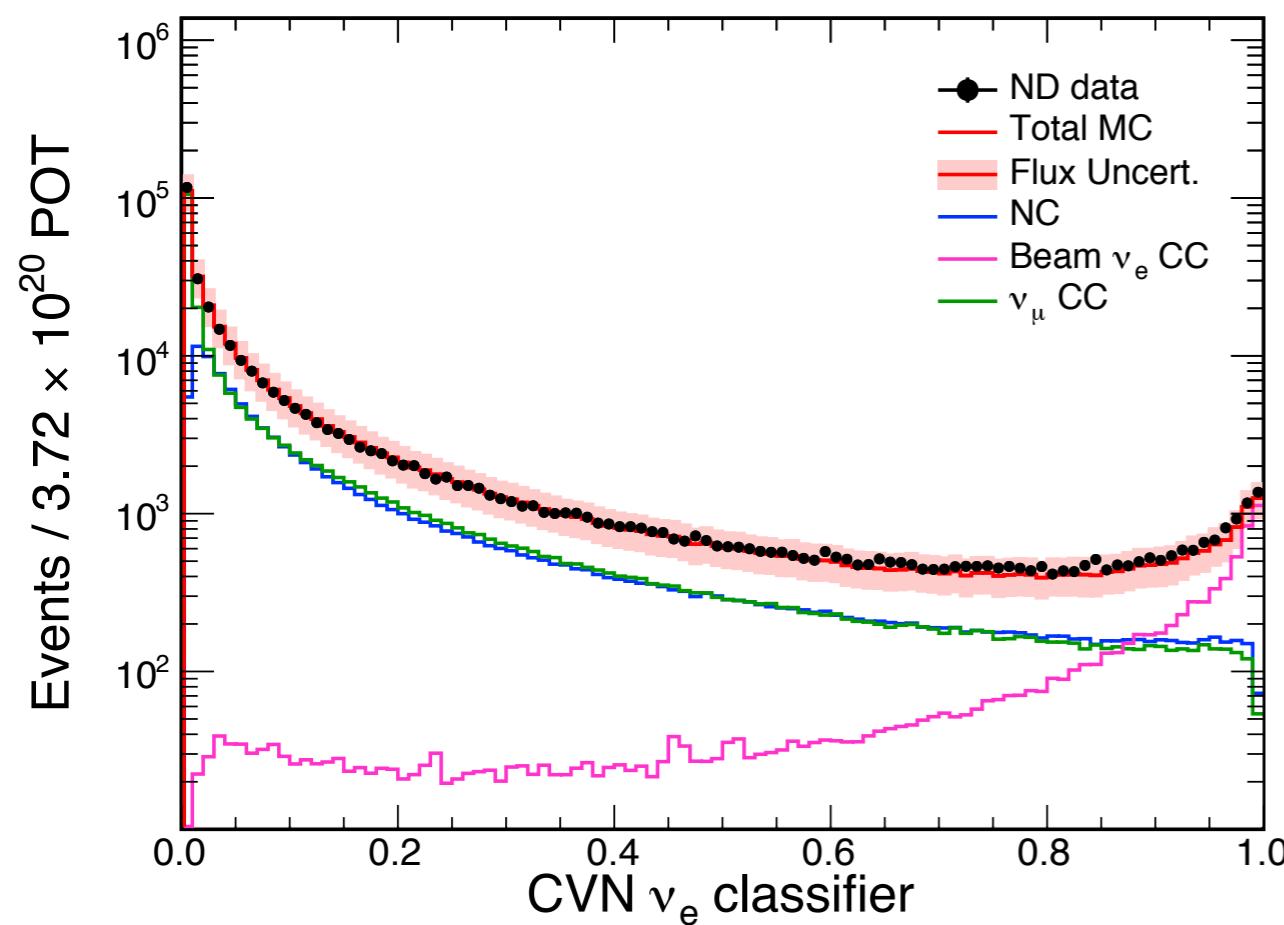


Performance on NearDet Data

NOvA Preliminary

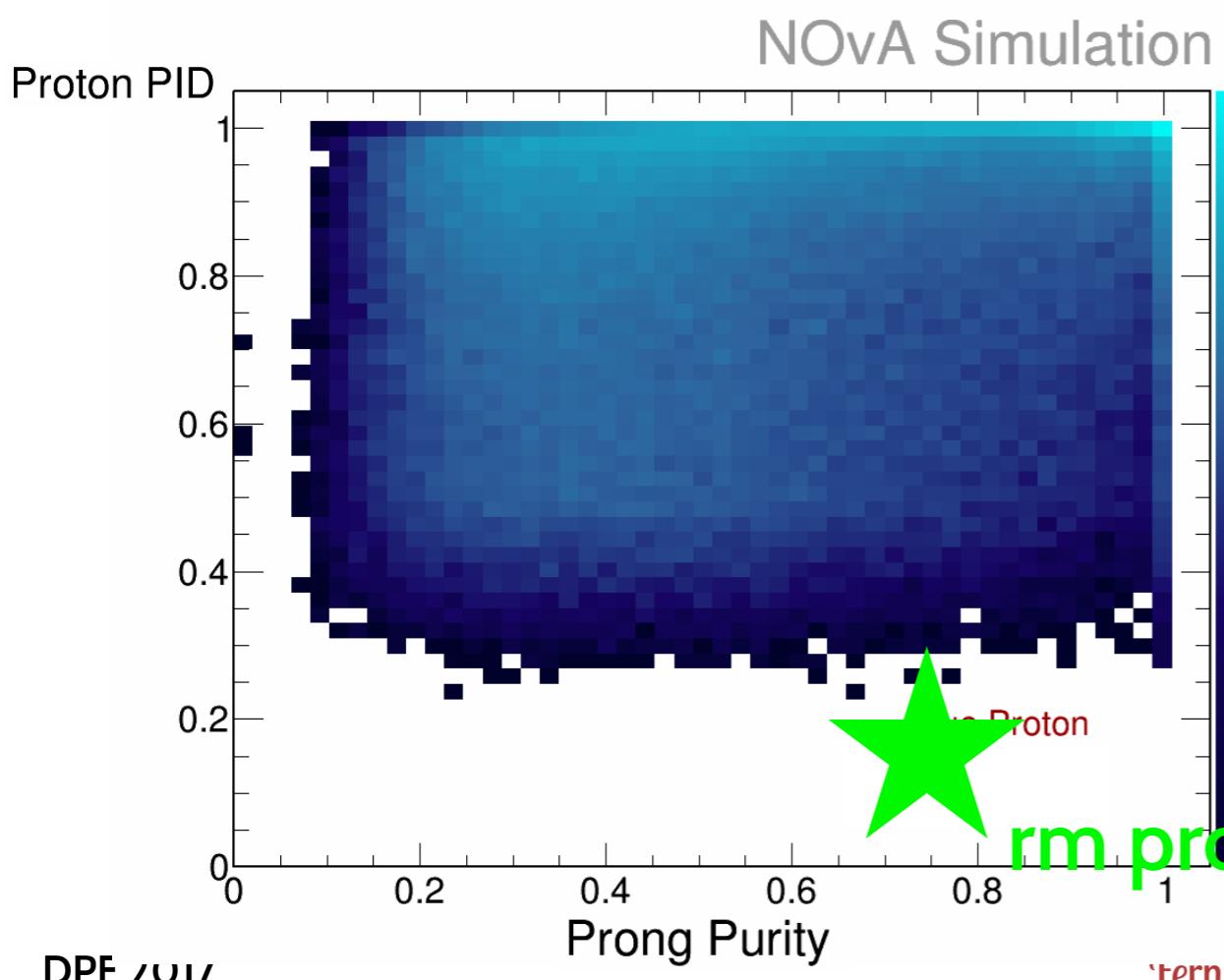
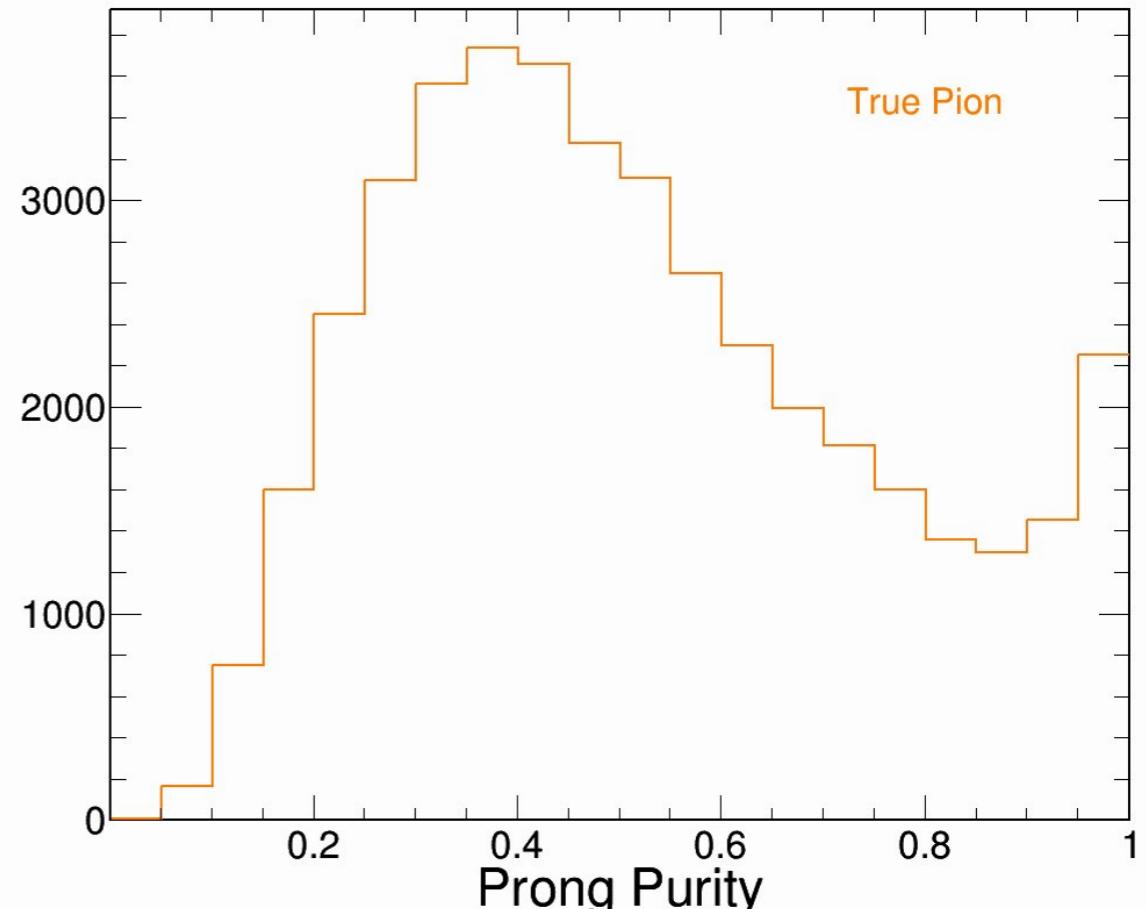


NOvA Preliminary

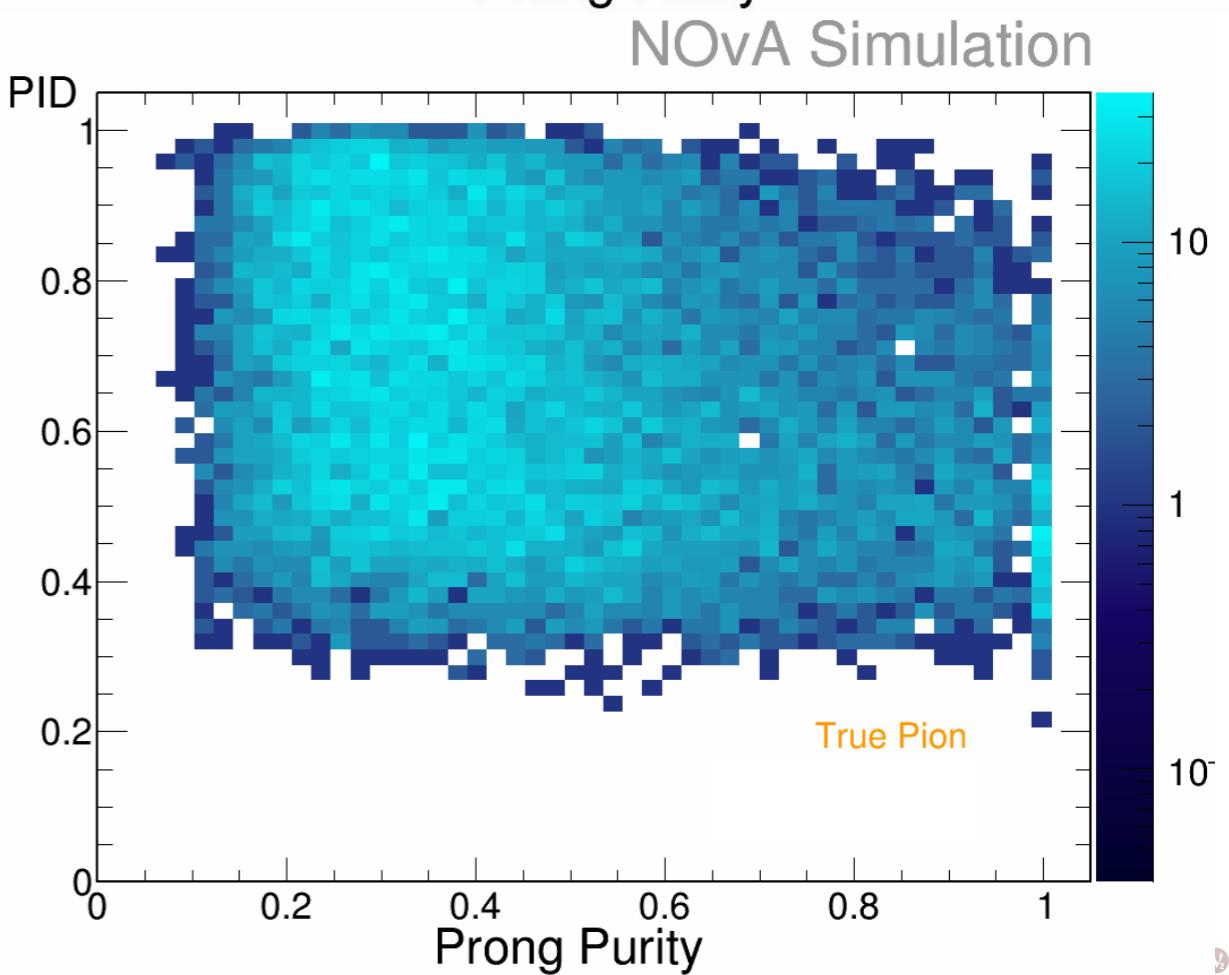


Caveats from the reconstruction

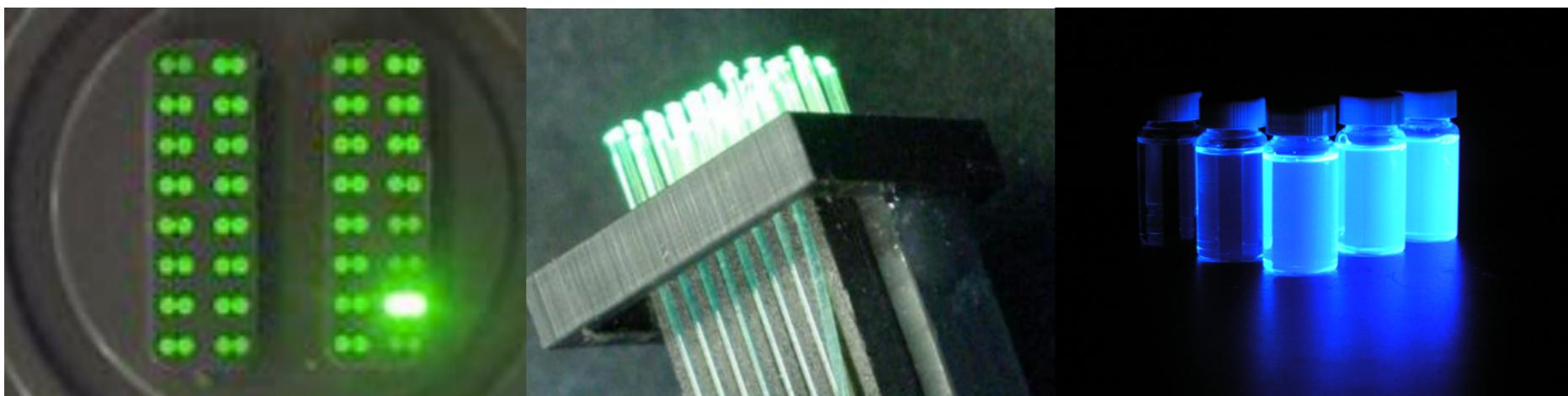
Prong quality impacts the performance of our classifier



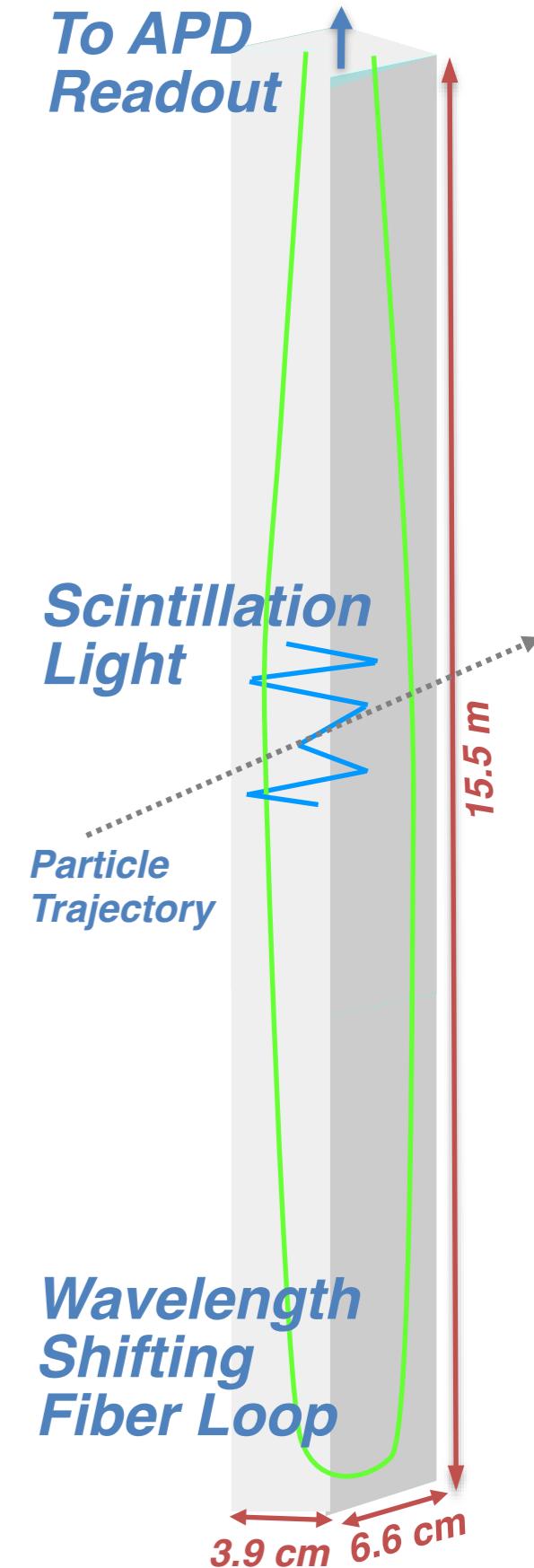
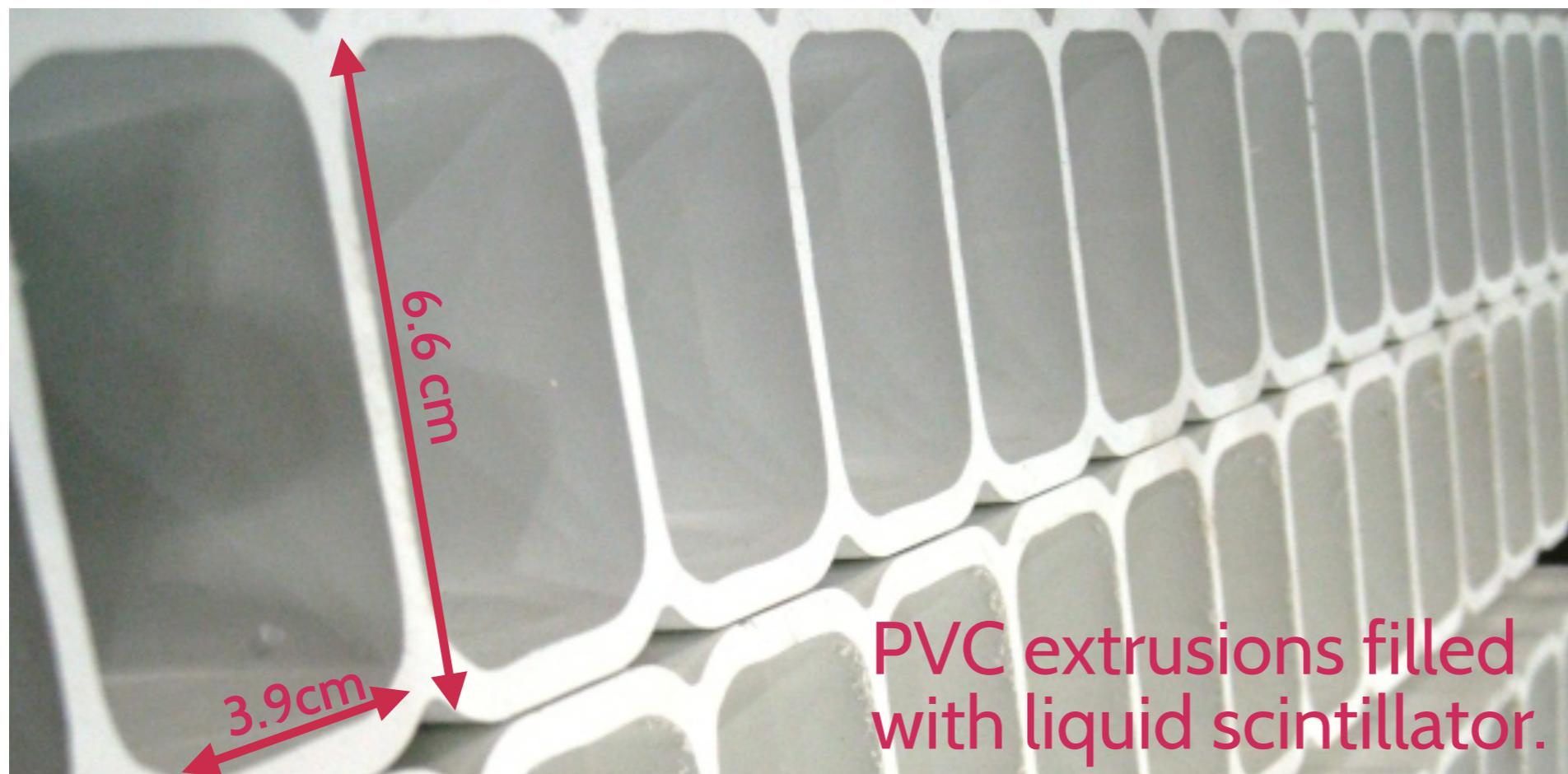
Terrianaa Psihas



The NO_νA Detectors

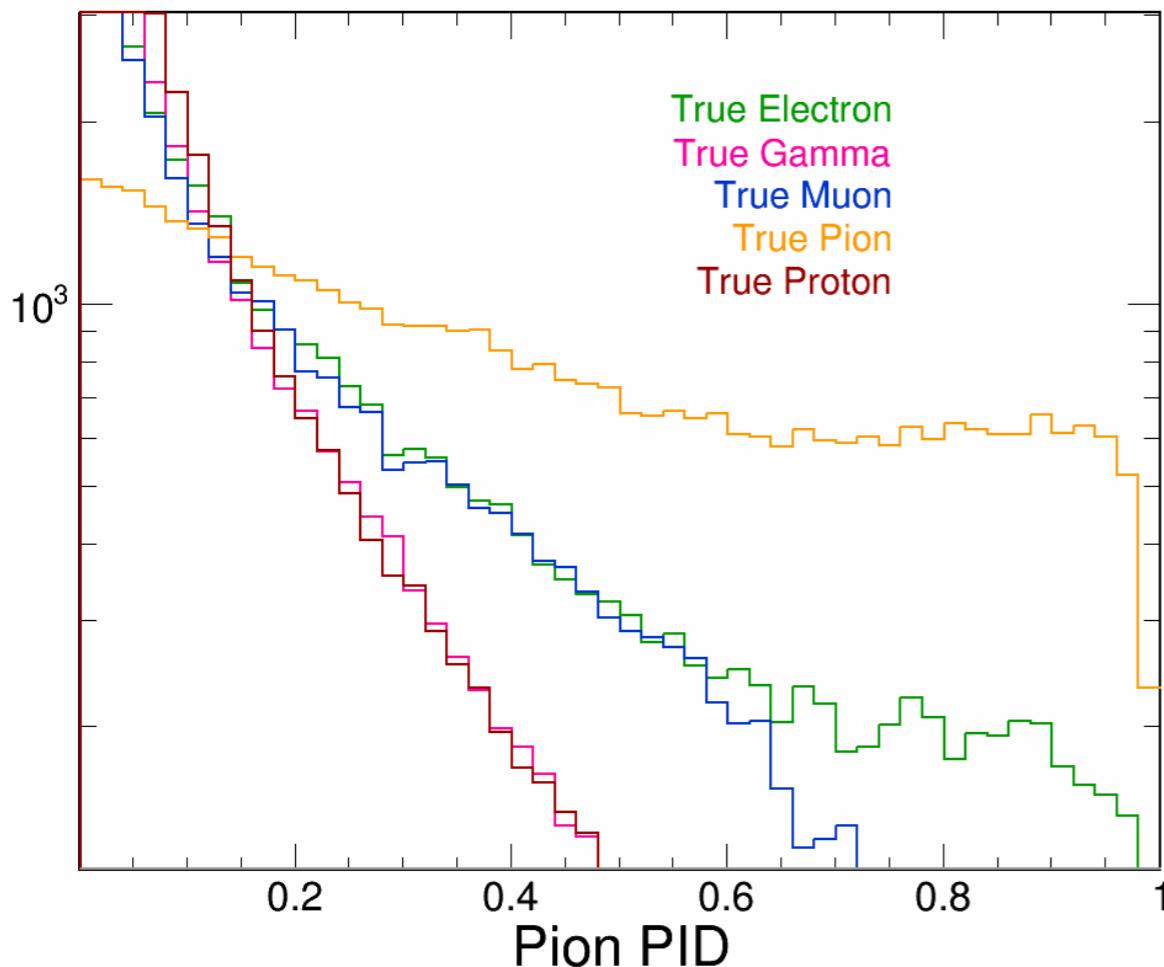


Charged particles are detected through the scintillation light produced in each cell.

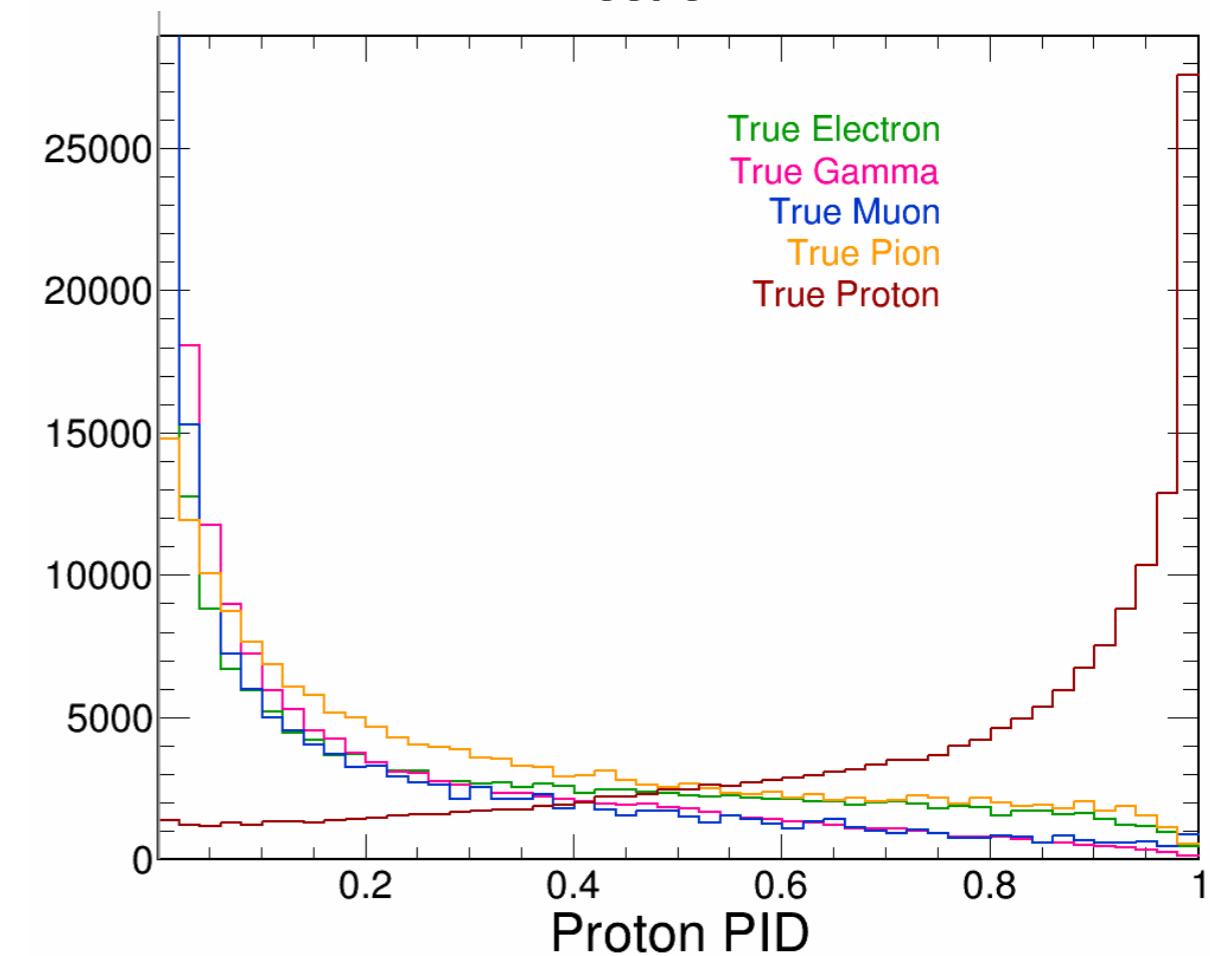
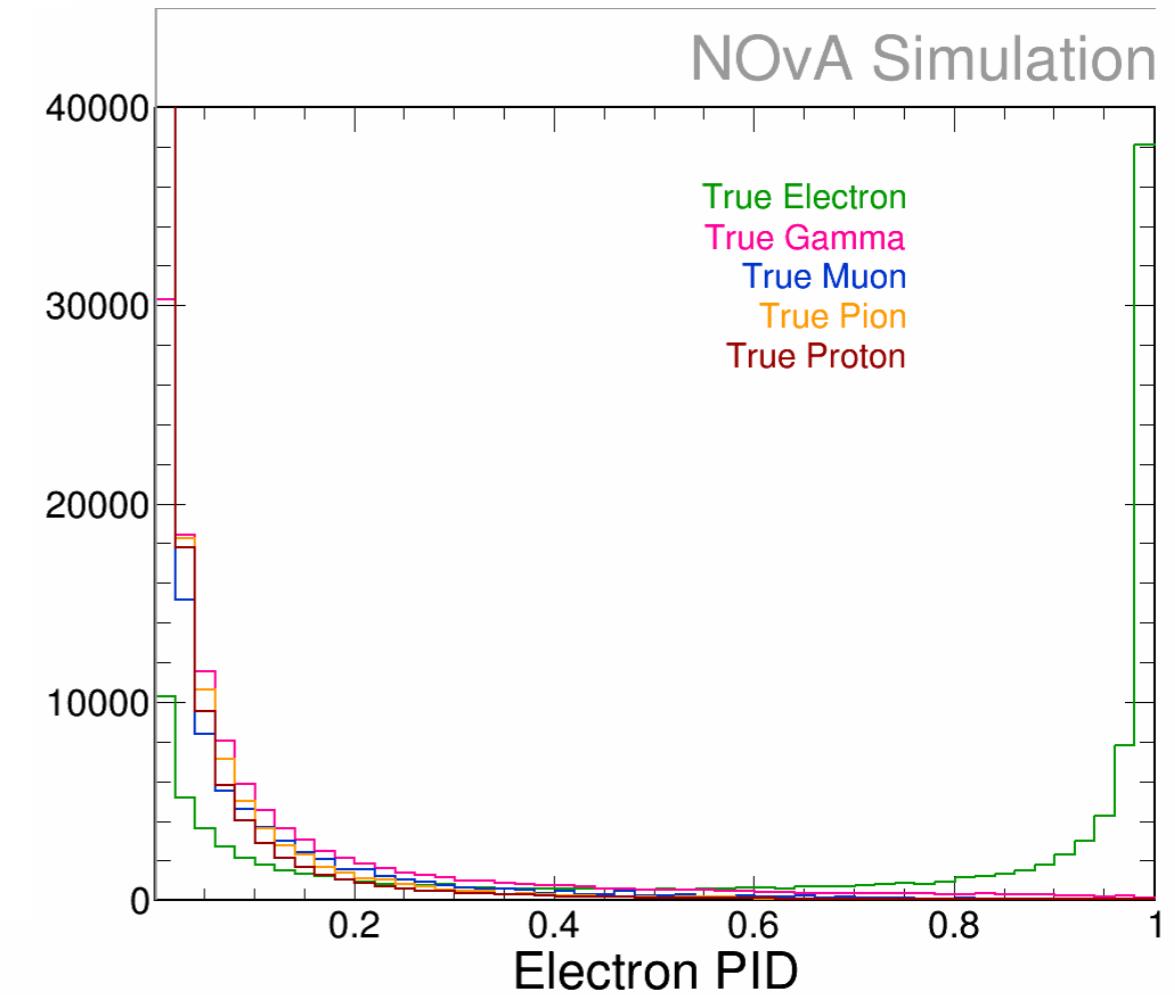


CVN prong results

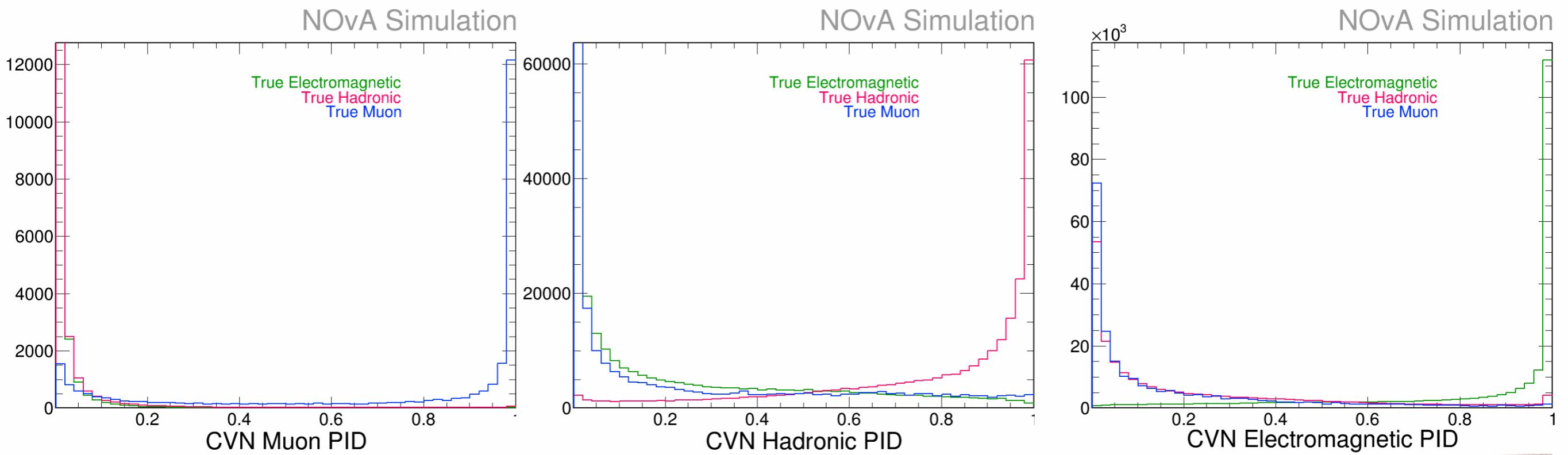
NOvA Simulation



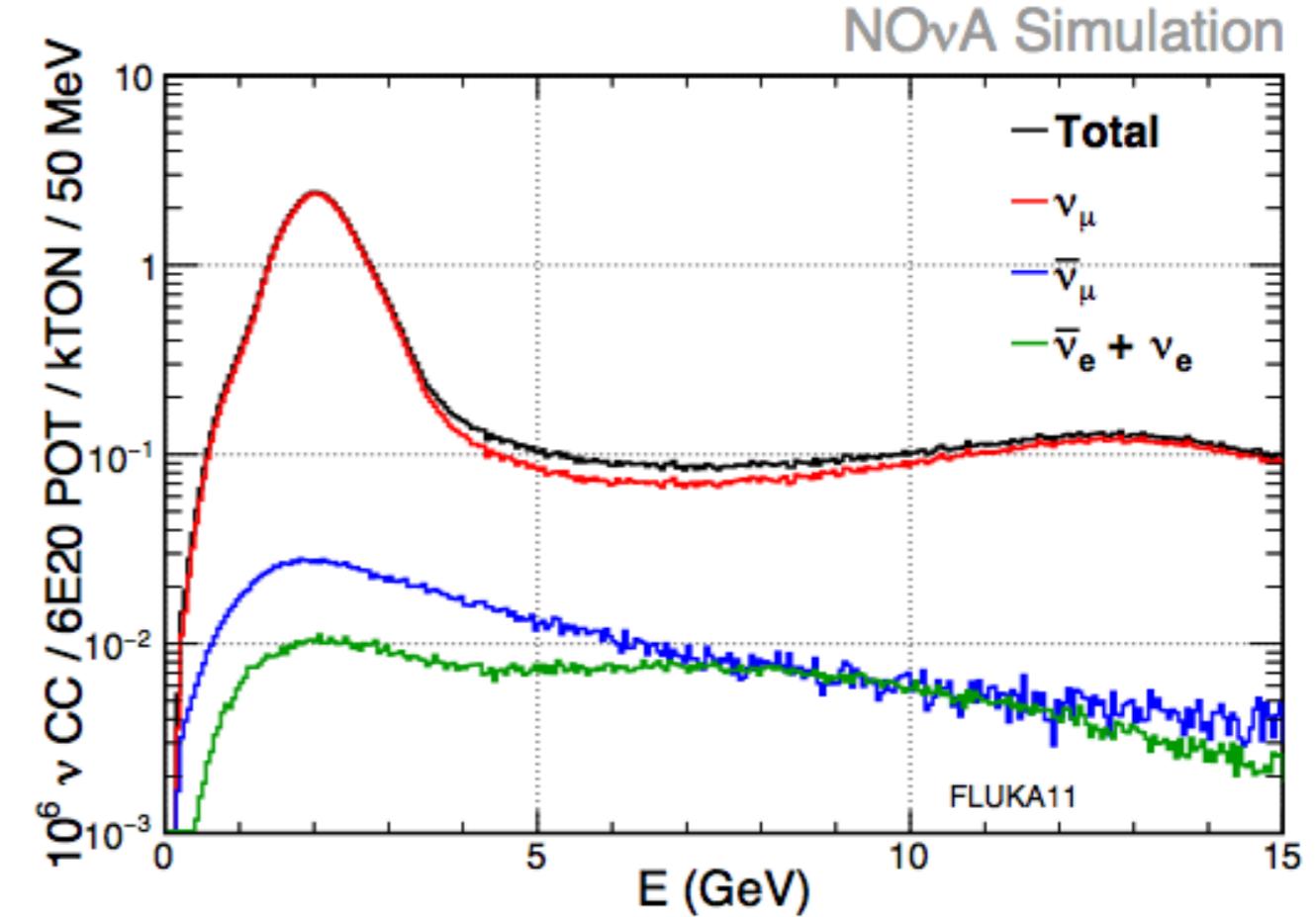
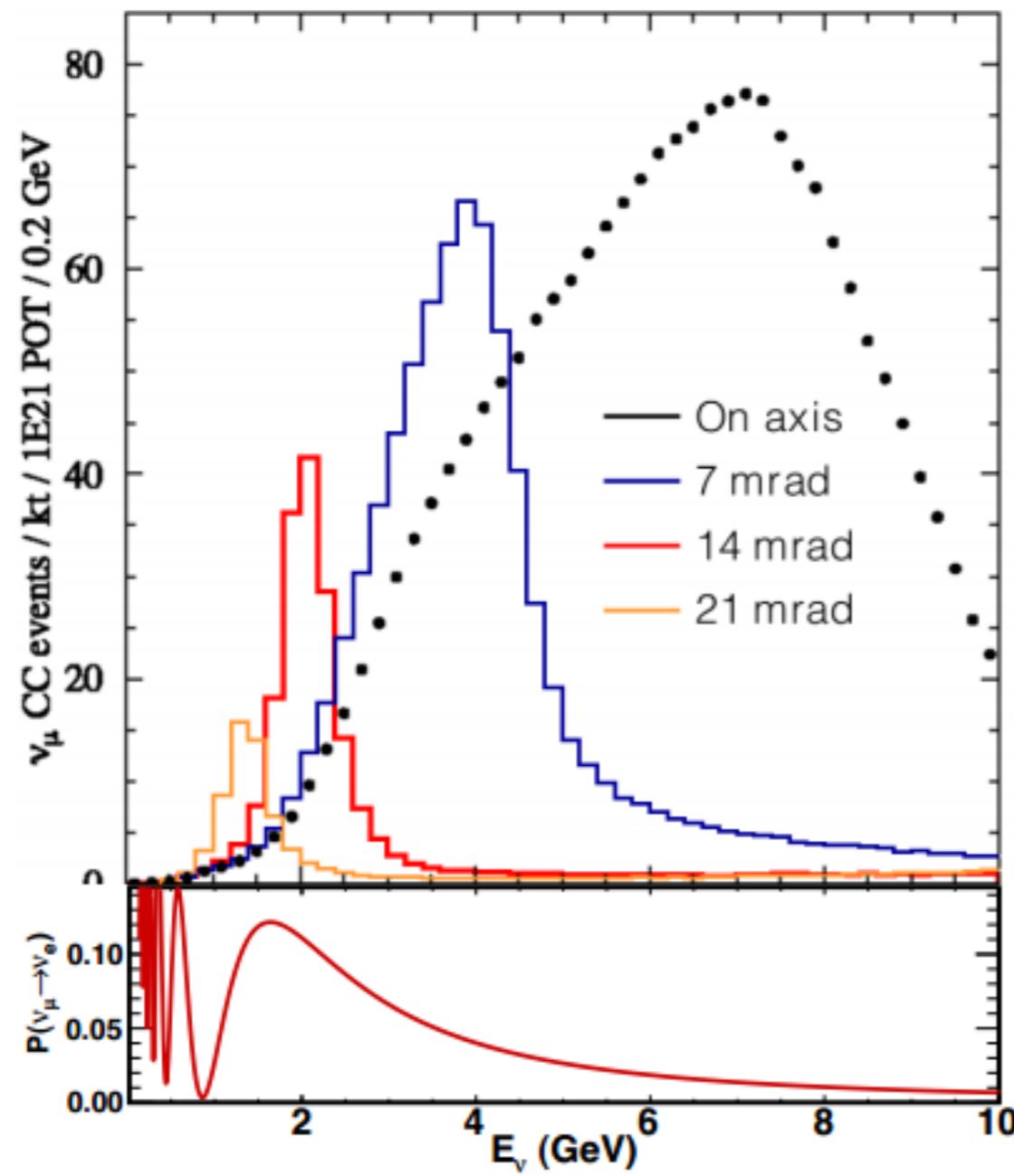
NOvA Simulation



CVN prong results



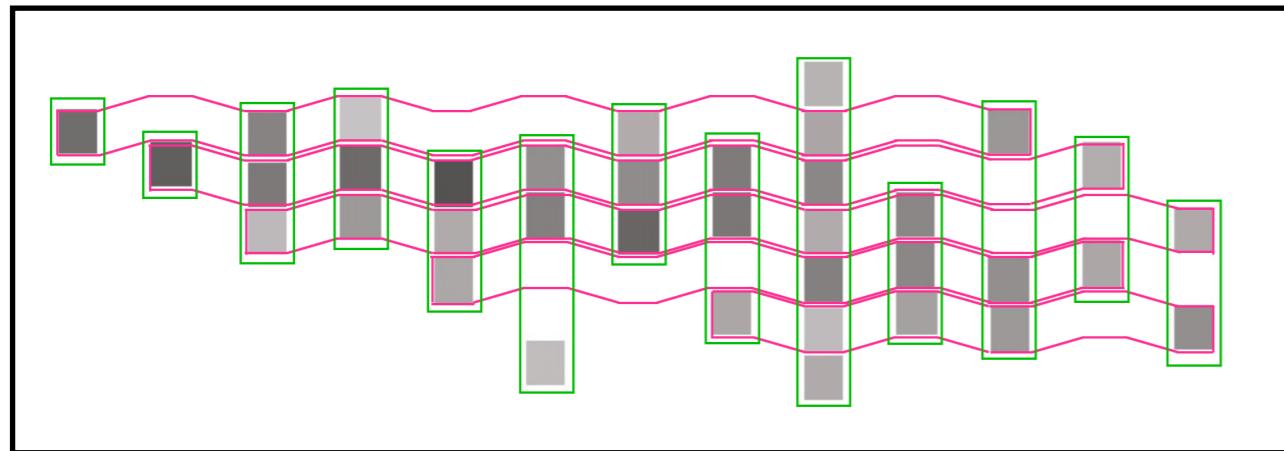
NuMI Beam



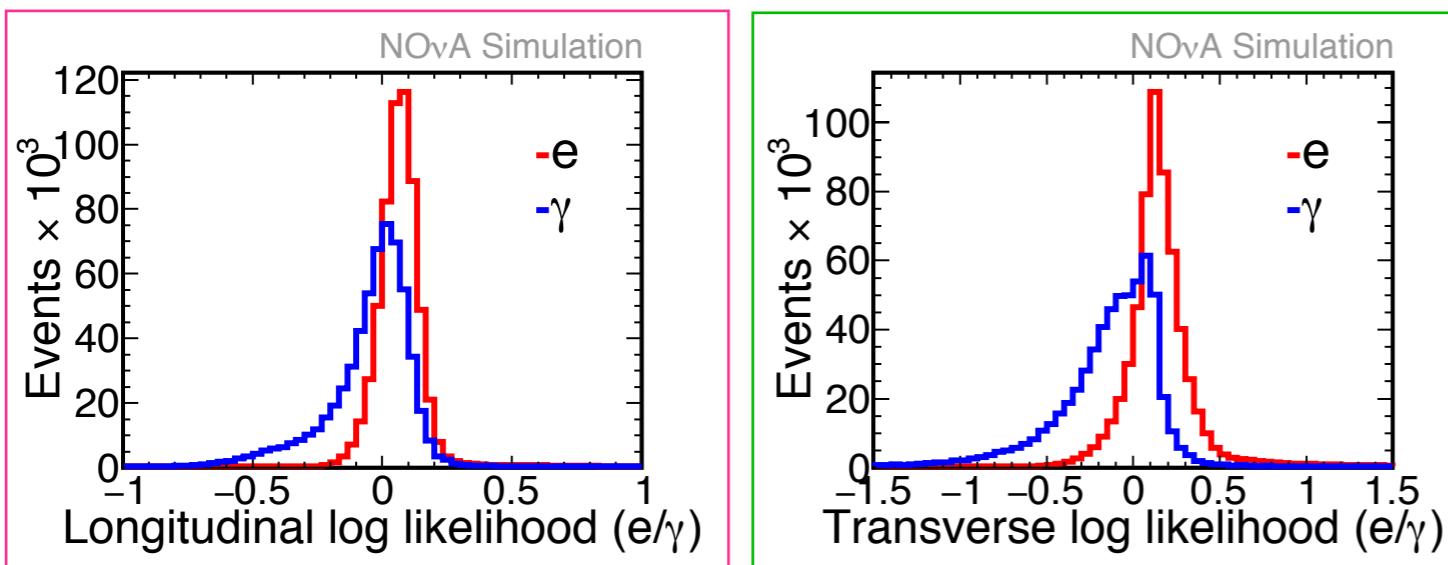
Traditional ID Methods

Mostly focused on identifying the lepton in the event. Extracted features (i.e. track length and scattering for muons, topology of energy depositions for electromagnetic showers)

- ★ **Require Previous reconstruction.**
- ★ **Features are pre-defined, based on MC or test data.**



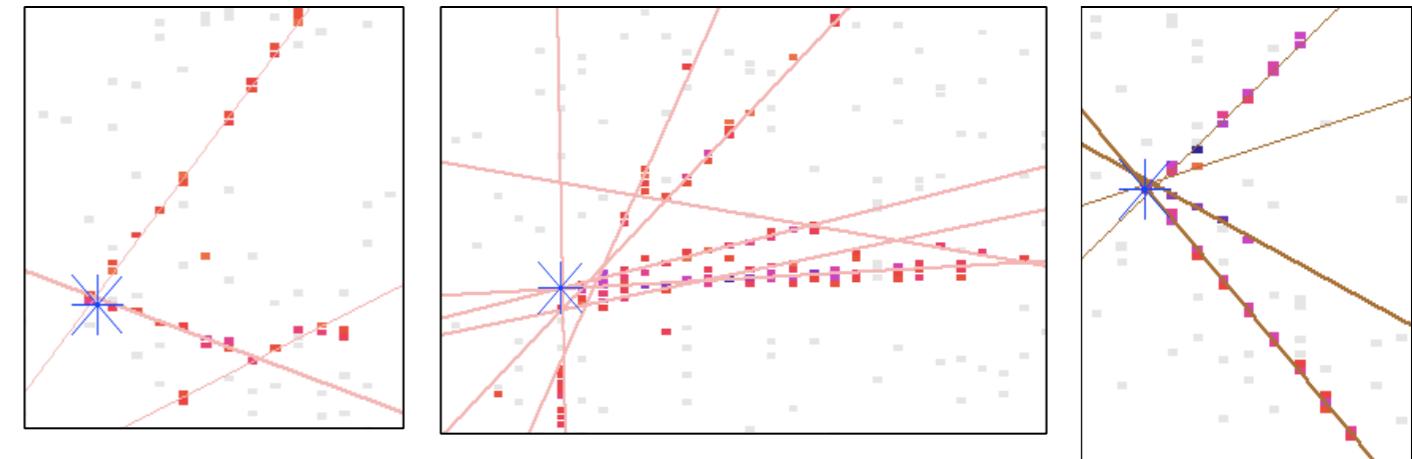
Example: The Likelihood ID method



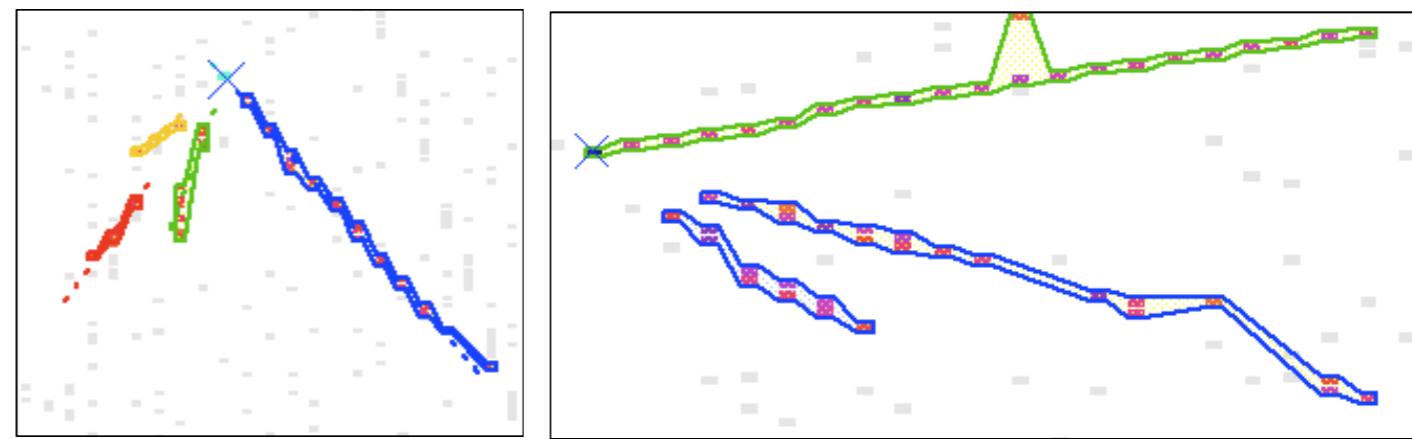
- ★ Reconstruct electron shower.
- ★ Find likelihoods from it's dE/dx profiles compared to particle hypotheses.

Likelihoods → Traditional Neural Network

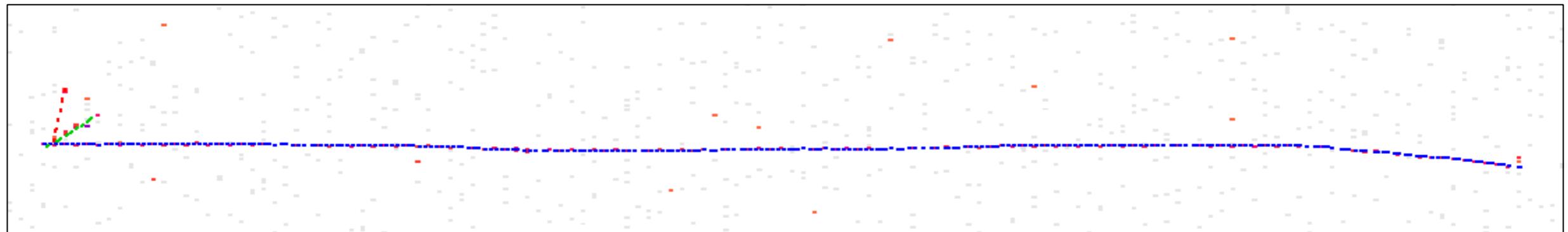
Vertexing: use lines of energy deposition formed with hough transforms to find intersections



Clustering: find clusters in angular space around the vertex and merge views via topology and prong dE/dx



Tracking: Trace particle trajectories using a kalman filter, example below

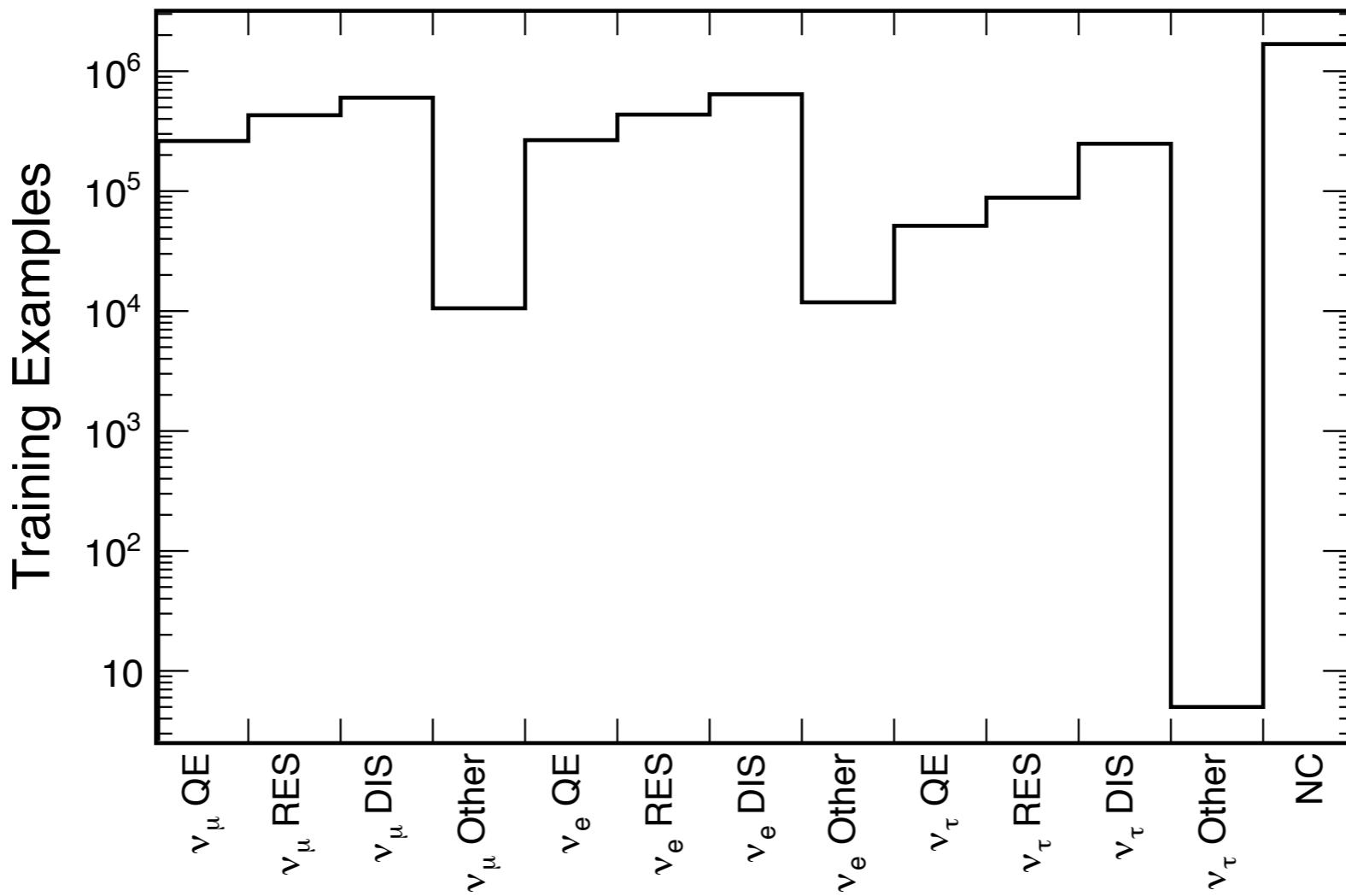


CVN for NOvA Events

Convolutional Visual Network

Neutrino Event CVN: *Siamese network architecture based on GoogLeNet.*

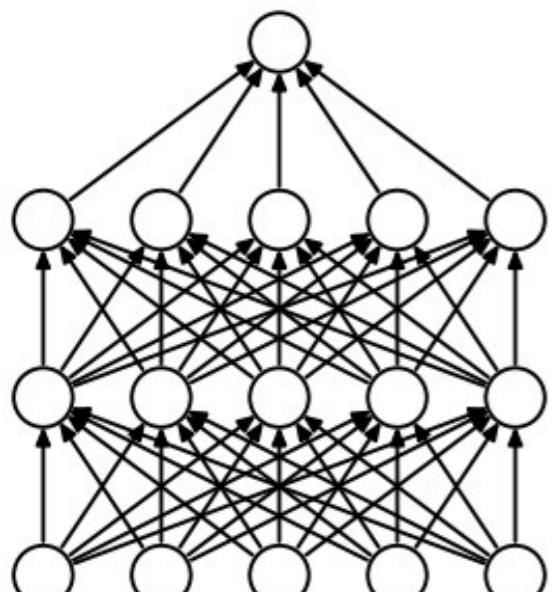
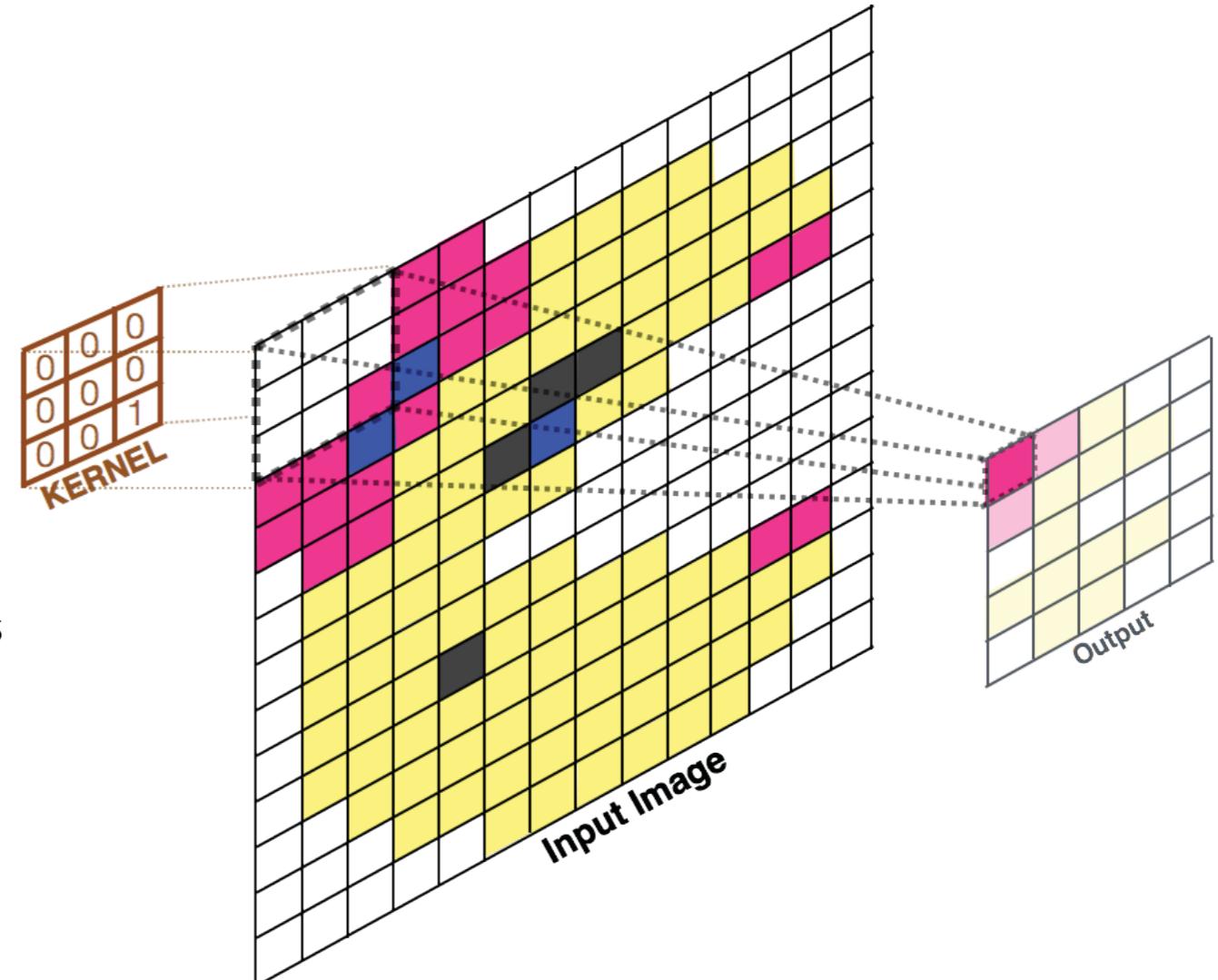
Training sample composition: We train on 4.7 million events of all interaction types plus cosmogenic backgrounds.



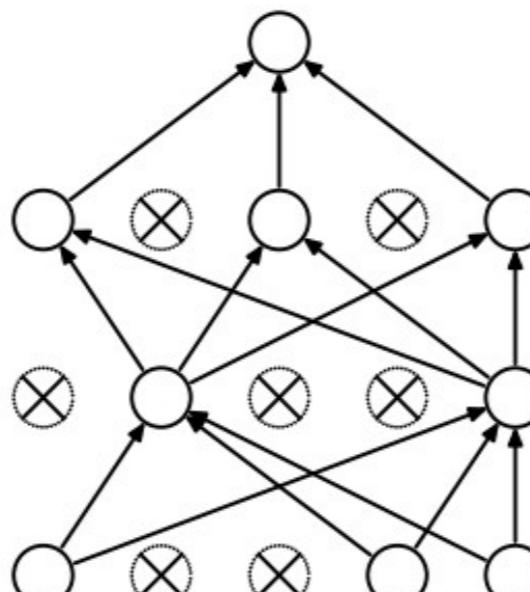
Network Layers

Kernel Renormalization:

Kernels evolve as the training progresses through renormalization. This process uses non saturating functions.



(a) Standard Neural Net



(b) After applying dropout.

Dropout:

Randomly reset weights, effectively removing whole nodes at each step.

Encourages complex dependence and discourages overtraining

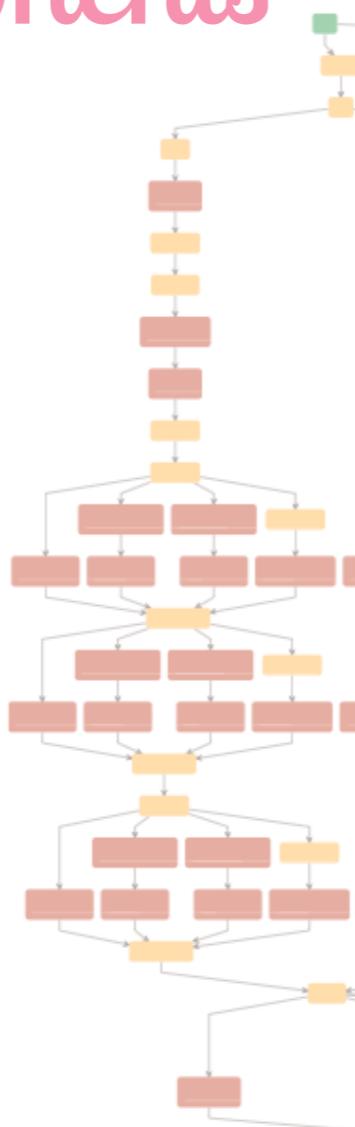
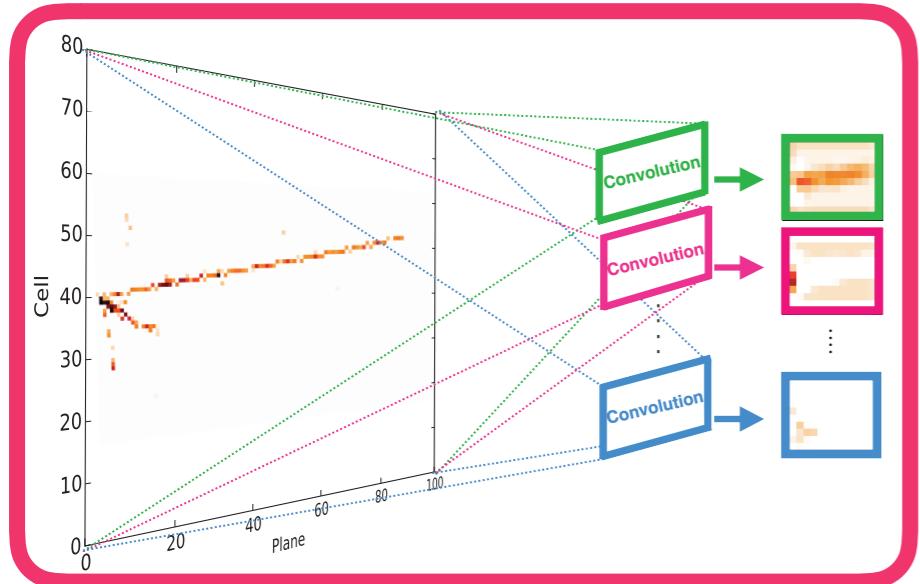


CVN Network Components

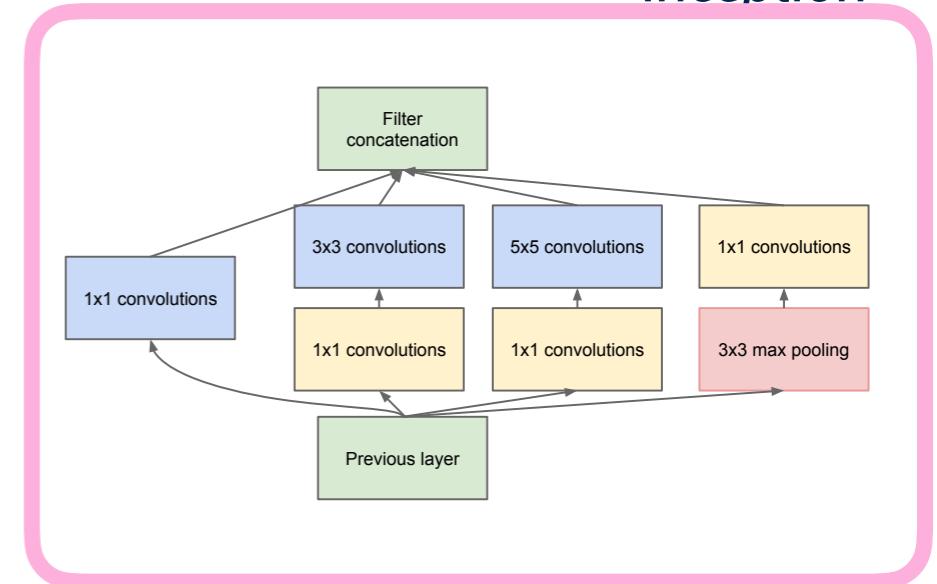
Convolutional Visual Network

Neutrino Event CVN: Siamese network architecture based on GoogLeNet.

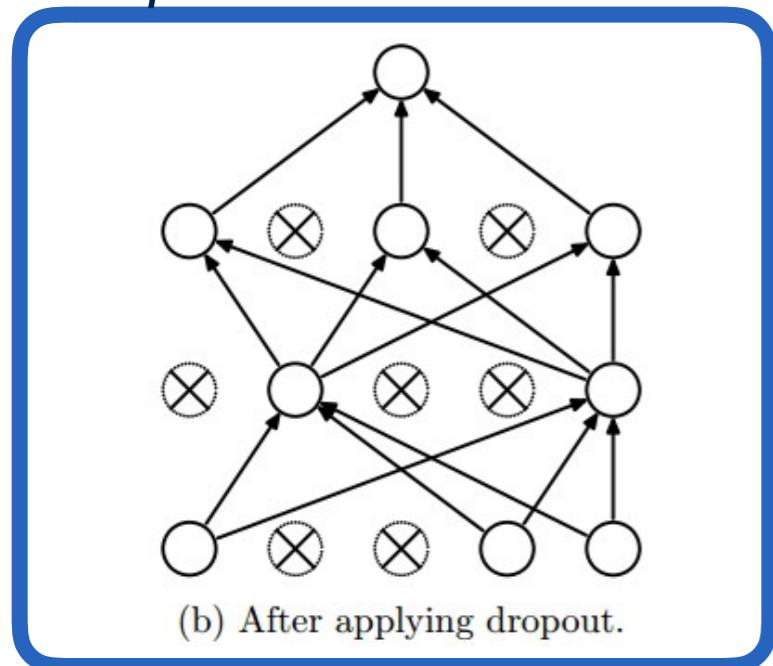
Convolution



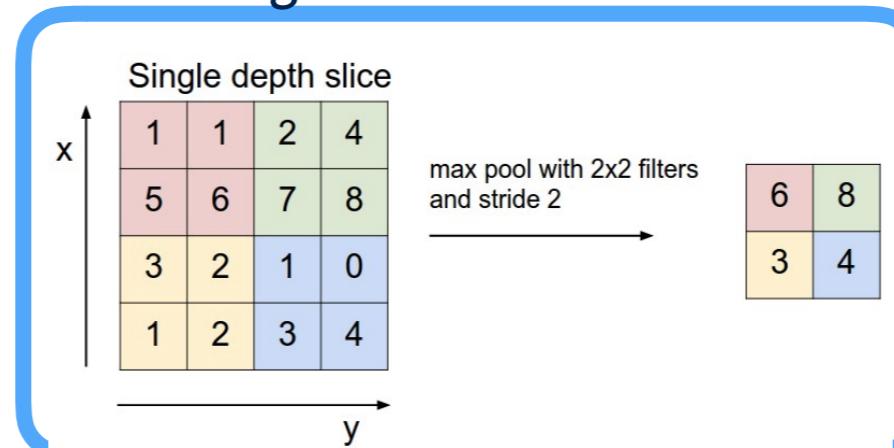
Inception



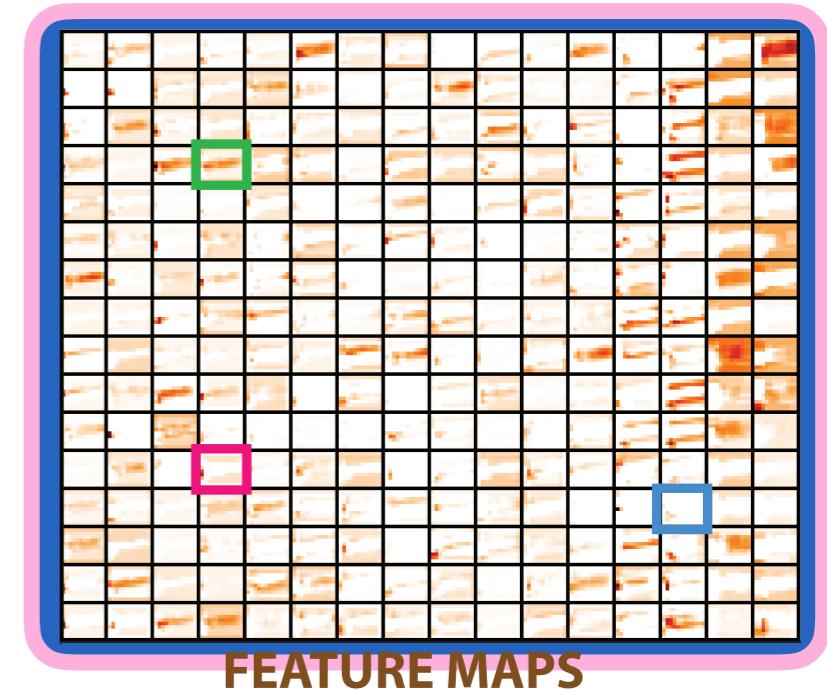
Dropout



Pooling

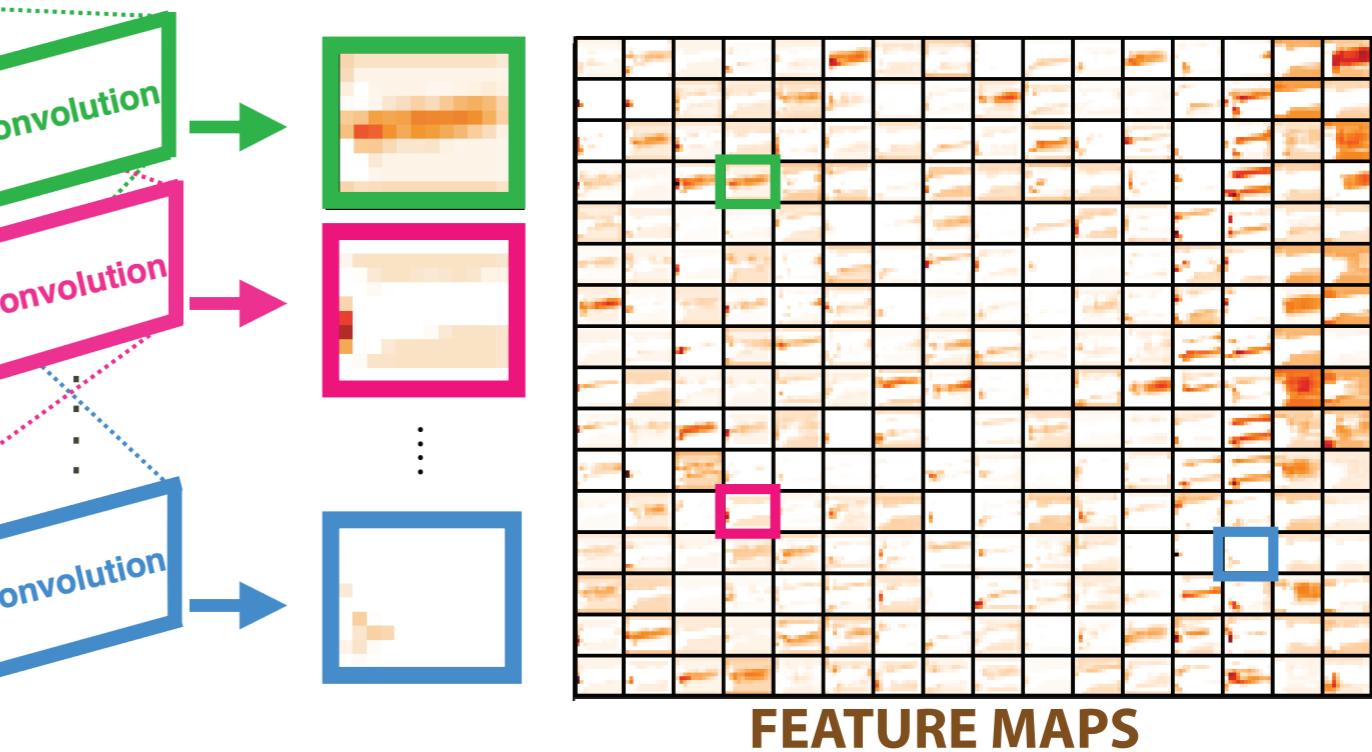
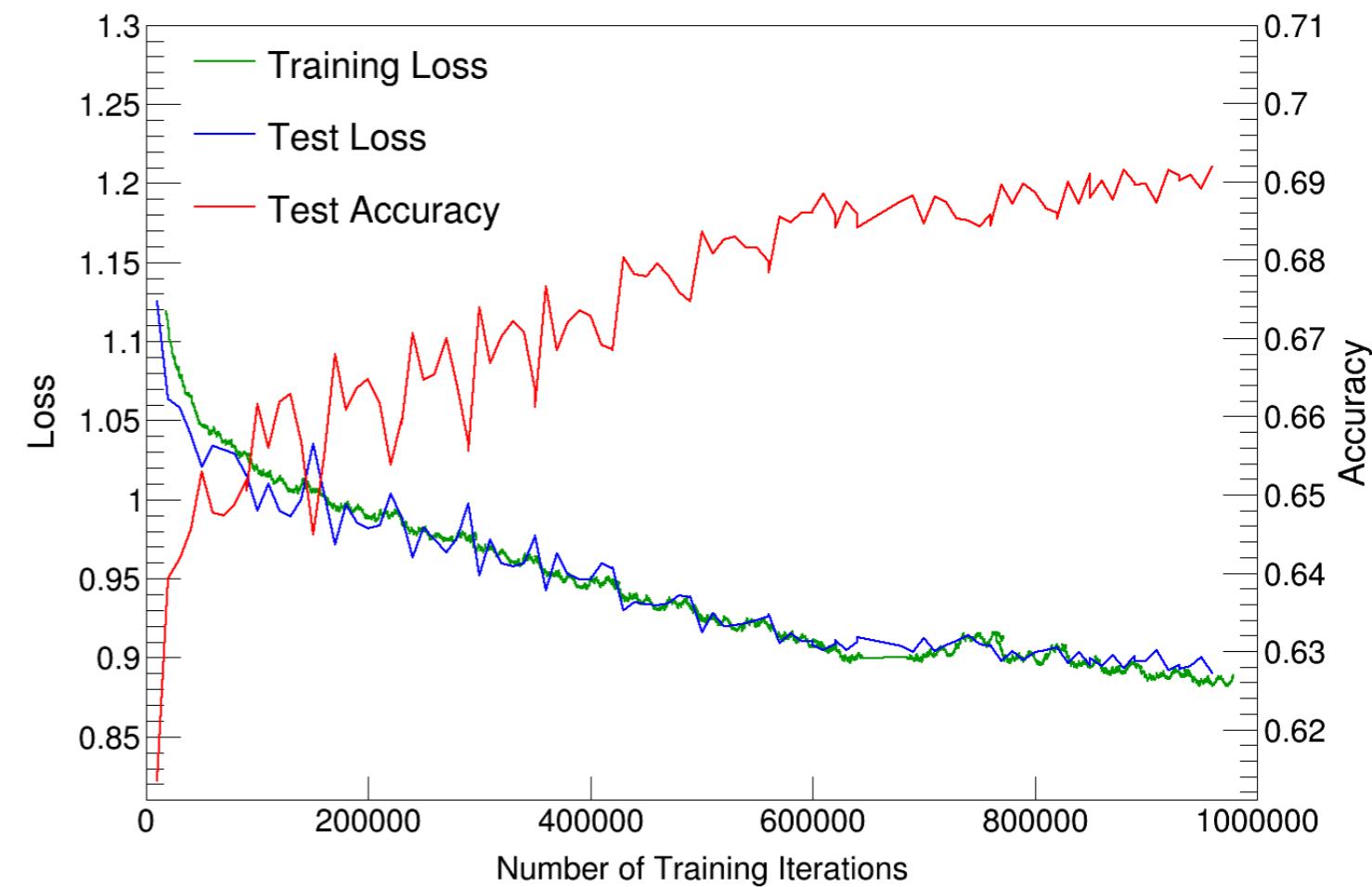
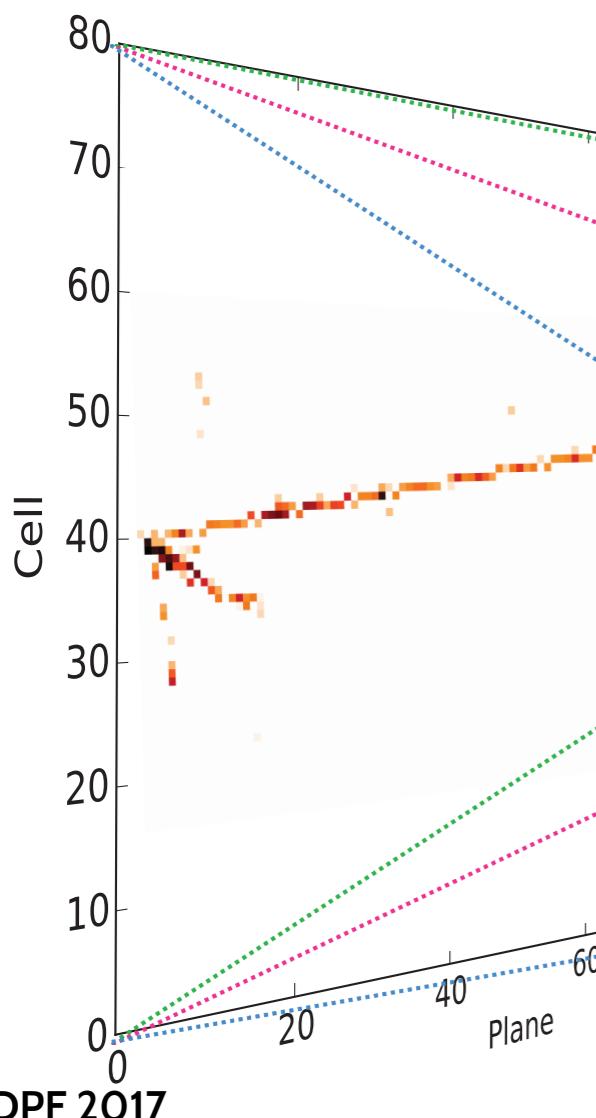


Inception output

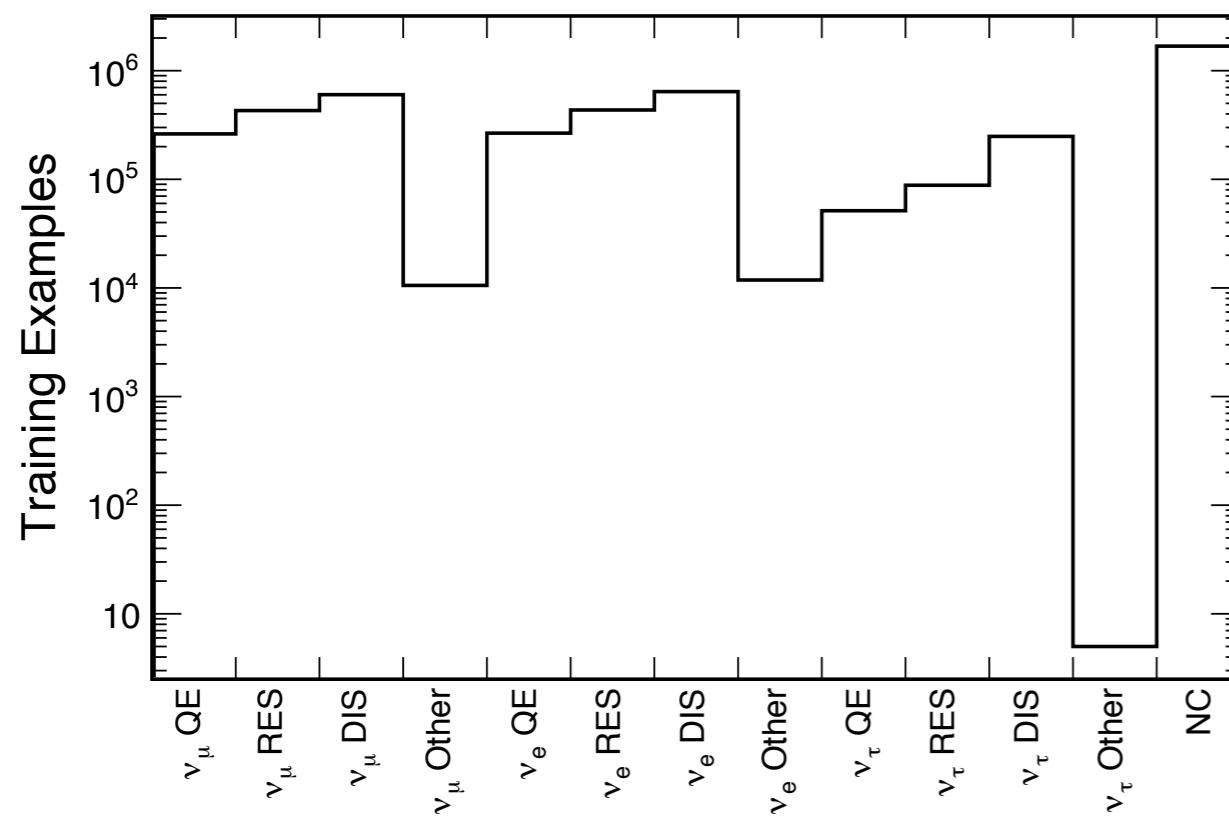


CVN Performance

- Trained on 4.7 million simulated events of all neutrino interaction types plus cosmic rays.
- training sample has minimal preselection

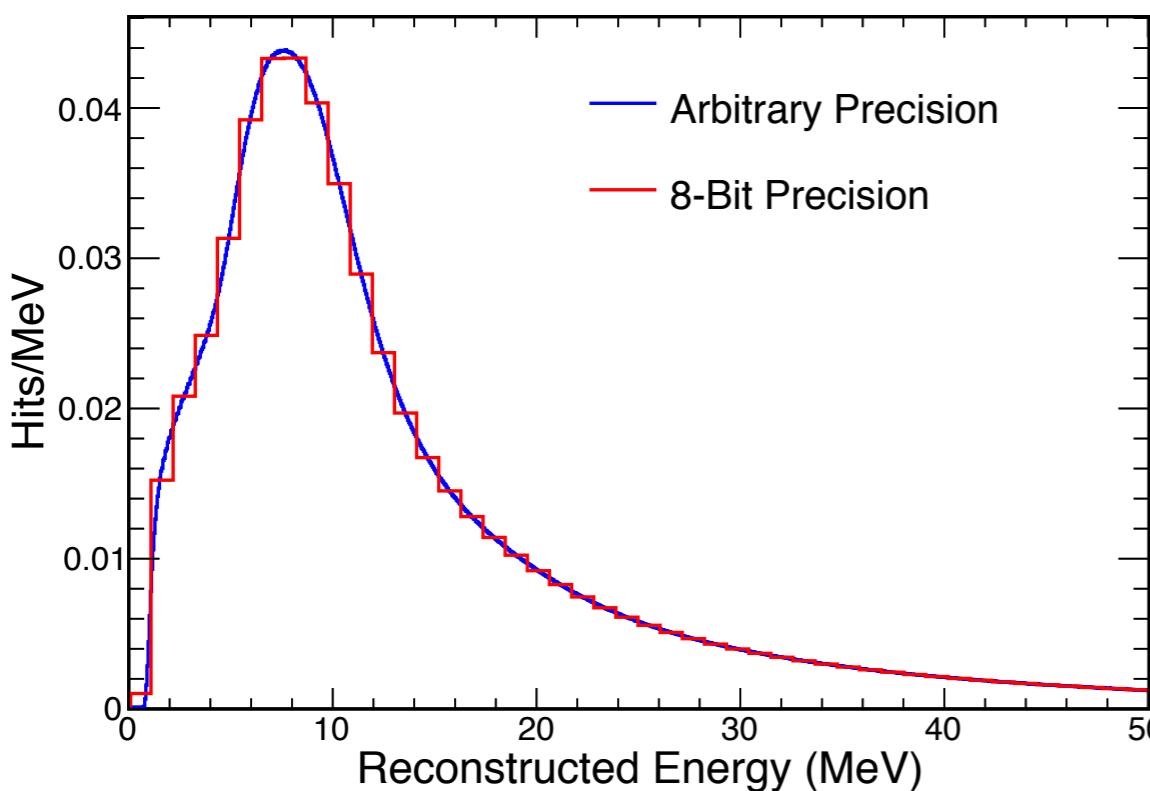


CVN Classifier



4.7 million, minimally preselected simulated events, pushed into LevelDB databases: 80% for training and 20% for testing.

Rescale calibrated energy depositions to go from 0 to 255 and truncate to chars for dramatically reduced file size at no loss of information

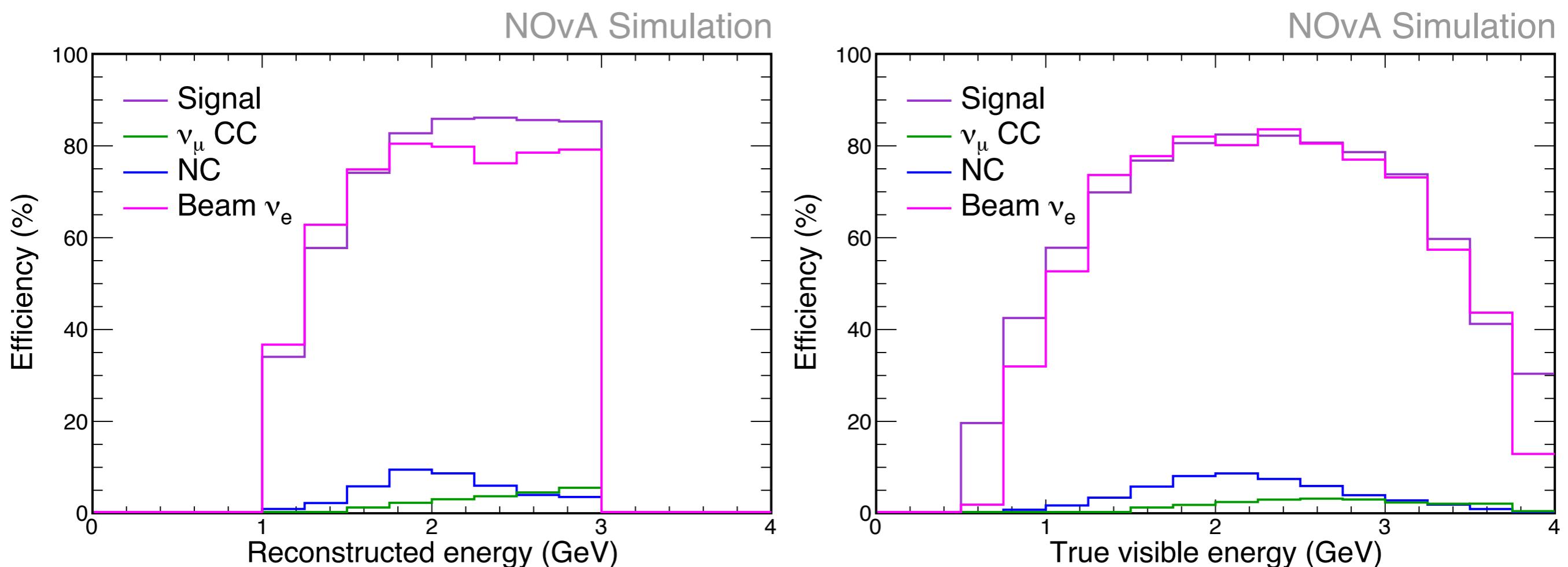


Fine tuned with 5 million cosmic data events taken from an out of beam time minimal bias trigger.

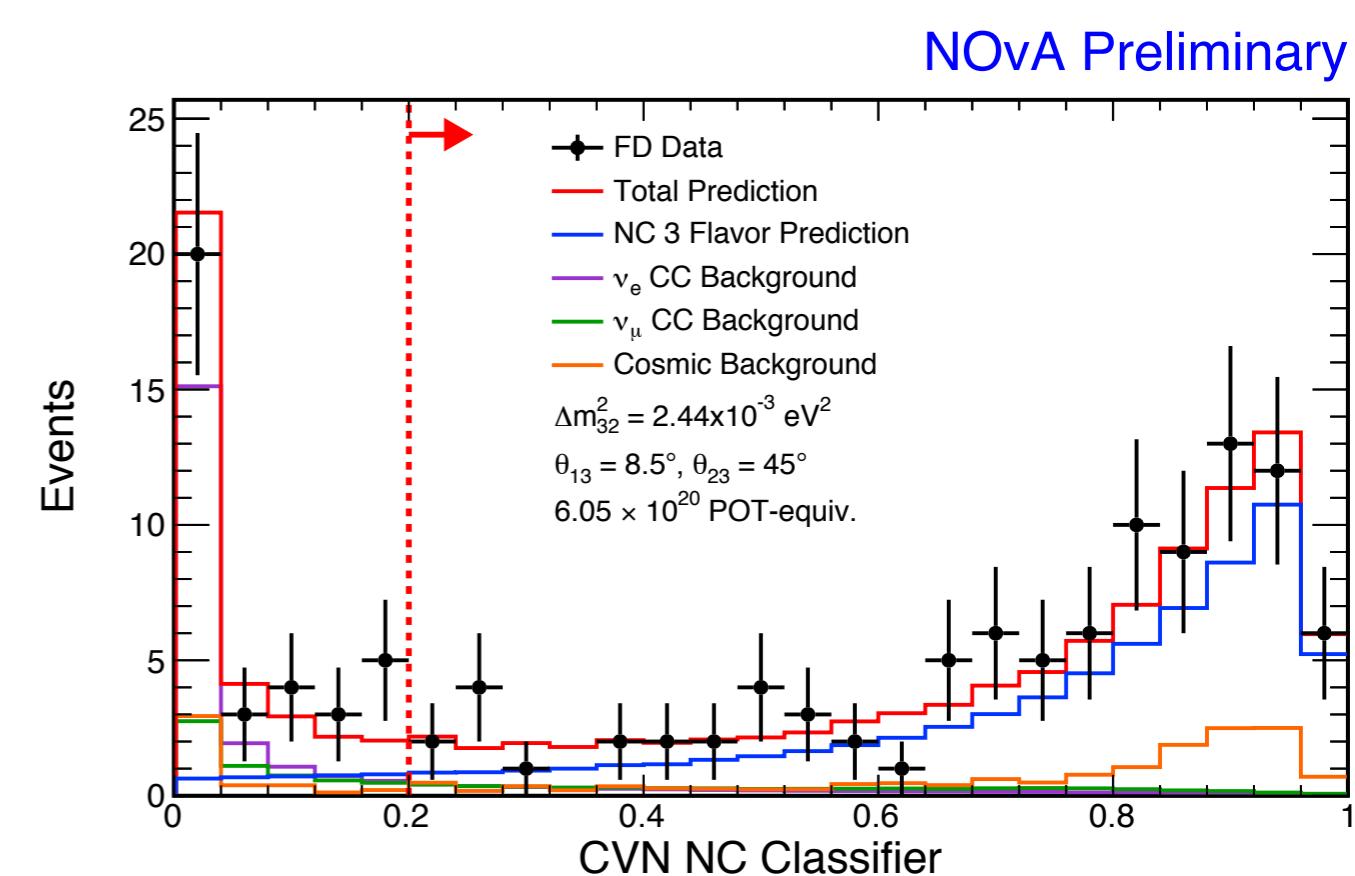
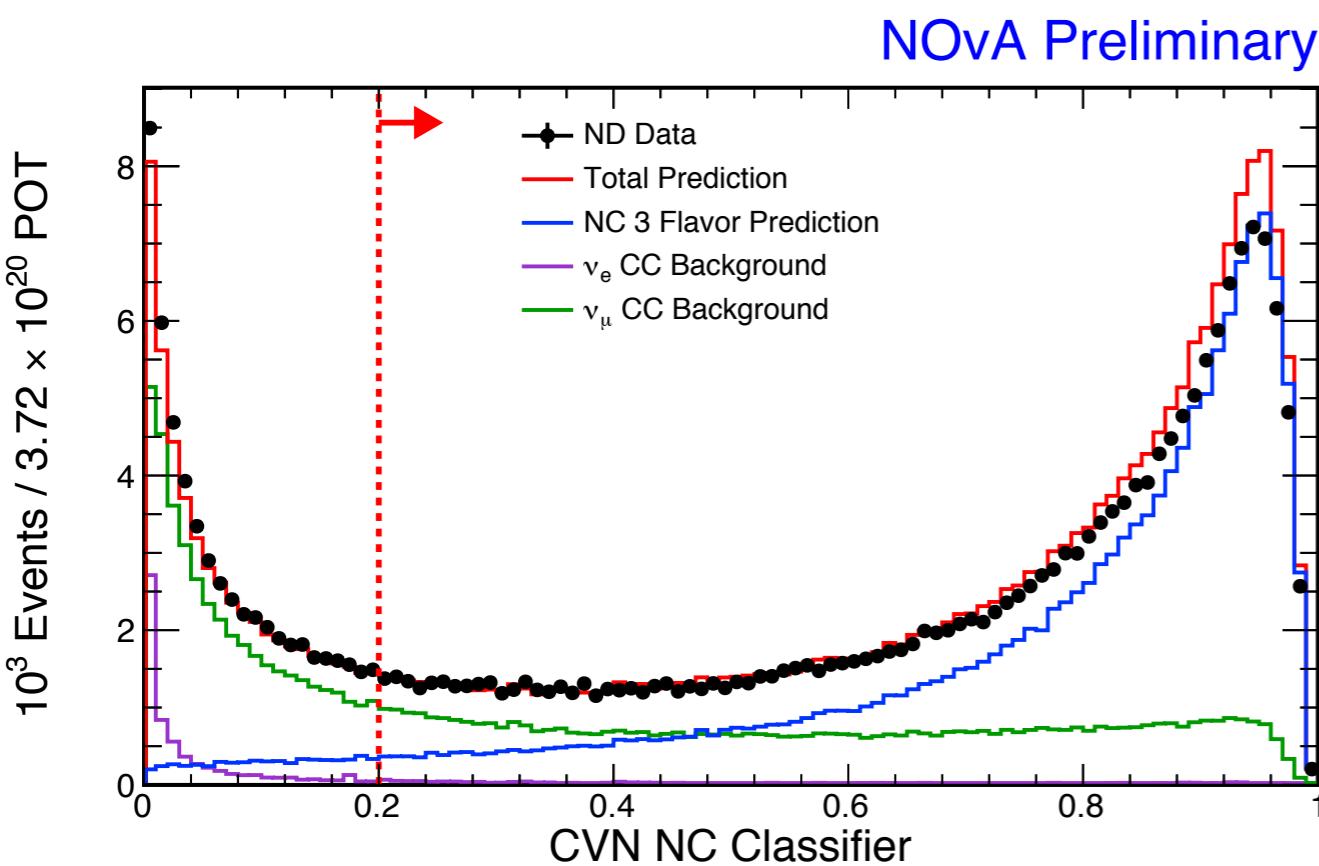
The architecture attempts to categorize events as $\{\nu_\mu, \nu_e, \nu_\tau\} \times \{\text{QE,RES,DIS}, \text{NC}\}$, or Cosmogenic.



CVN MC Efficiency



Neutral Current Neutrino Analysis

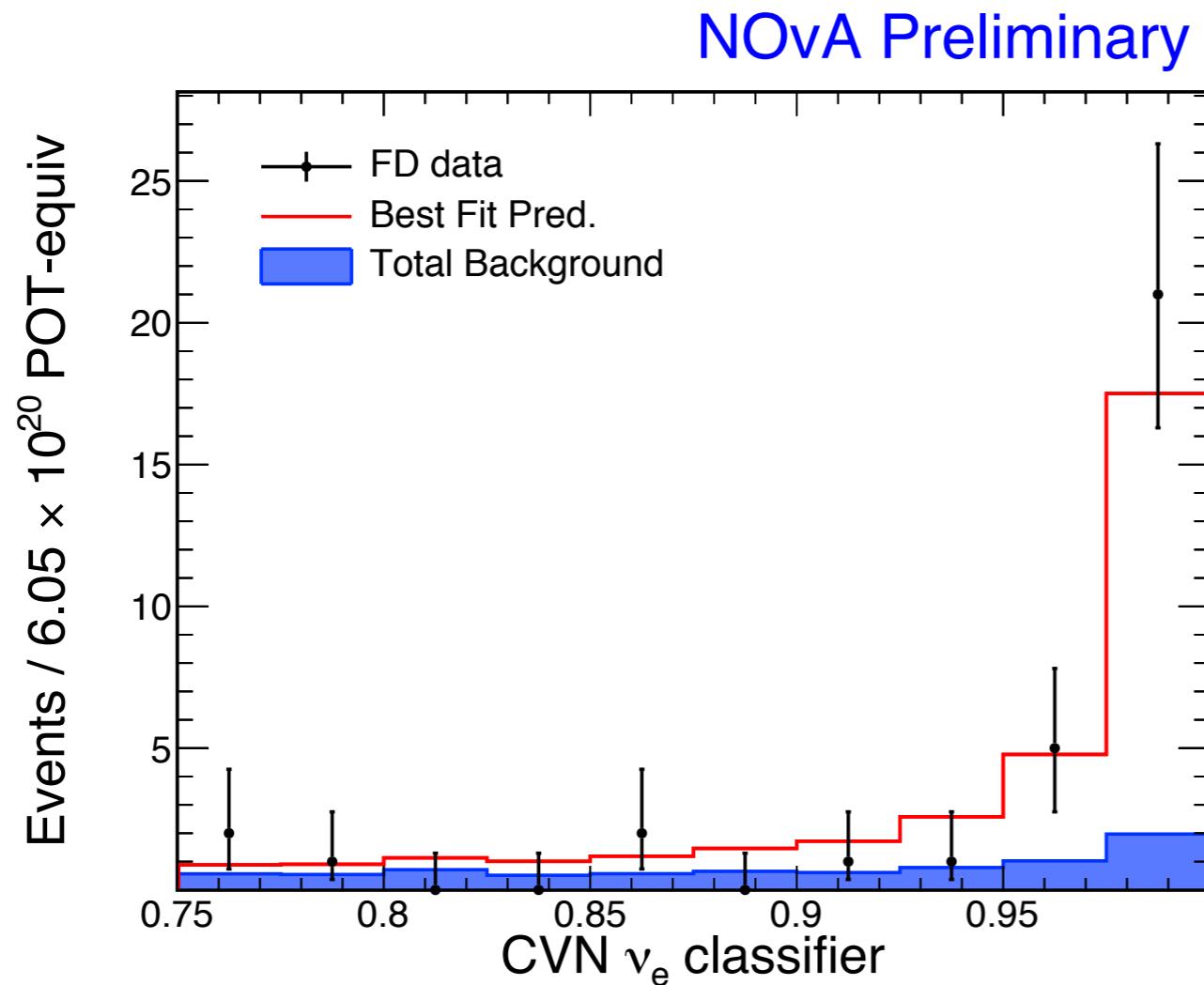


CVN Performance on ν_e

Implemented in NOvA's main analysis for the results shown this summer at Neutrino 2016 was the **first implementation of a CNN in a HEP result.**

Total bkg	NC	Beam ν_e	ν_μ CC	ν_τ CC	Cosmogenic
8.2	3.7	3.1	0.7	0.1	0.5

33 events selected with estimated background of ~8



76% Purity
73% Efficiency

An equivalent increased exposure of 30%

